CHAPTER 3

USING PROBABILISTIC ANALYSIS IN HUMAN HEALTH ASSESSMENT

3.0 INTRODUCTION

This chapter outlines how probabilistic analysis may be applied to human health risk assessments in the Environmental Protection Agency's (EPA) Superfund program. The paradigm for human health risk assessment as described in EPA's *Risk Assessment Guidance for Superfund* (U.S. EPA, 1989), includes data collection/evaluation in addition to exposure and toxicity assessment and risk characterization. Although the strategies and methods used in collecting and analyzing data can significantly impact the uncertainty in a risk estimate, they are issues relevant to risk assessment in general, and are addressed in other guidance documents, such as EPA's *Guidance for Data Useability in Risk Assessment* (U.S. EPA, 1992b). RAGS Volume 3: Part A focuses on a tiered approach for incorporating quantitative information on variability and uncertainty into risk management decisions.

3.1 CHARACTERIZING VARIABILITY IN EXPOSURE VARIABLES

Exhibit 3-1 gives the general equation used for calculating exposure, often expressed as an average daily intake. In a point estimate approach, single values (typically a mixture of average and high-end values) are input into the equation. In probabilistic risk assessment (PRA), the only difference is that a probability distribution, rather than single value, is specified for one or more variables. A Monte Carlo simulation is executed by repeatedly selecting random values from each of these distributions and calculating the corresponding exposure and risk. For the majority of PRAs, it is expected that probability distributions will be used to characterize inter-individual variability, which refers to true heterogeneity or diversity in a population. Thus, variability in daily intake, for example, can be characterized by combining multiple sources of variability in exposure, such as ingestion rate, exposure frequency, exposure duration, and body weight. Variability in chemical concentrations (Chapter 5 and Appendix C) and the toxicity term in ecological risk assessment (Chapter 4) may also be considered in risk calculations.

EXHIBIT 3-1
GENERAL EQUATION FOR EXPOSURE

$$I = \frac{C \times CR \times EF \times ED}{BW \times AT} \qquad \text{Eq. 3-1}$$

where,

I	=	daily intake
C	=	contaminant concentration
CR	=	contact rate (ingestion, inhalation, dermal contact)
EF	=	exposure frequency
ED	=	exposure duration
BW	=	body weight
AT	=	averaging time

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3

95% UCL for mean - The one-sided 95% upper confidence limit for a population mean; if a sample of size (*n*) was repeatedly drawn from the population, the 95% UCL will equal or exceed the true population mean 95% of the time. It is a measure of uncertainty in the mean, not to be confused with the 95th percentile (see below), which is a measure of variability. As sample size increases, the difference between the UCL for the mean and the true mean decreases, while the 95th percentile of the distribution remains relatively unchanged.

95th percentile - The number in a distribution that is greater than 95% of the other values of the distribution, and less than 5% of the values. When estimated from a sample, this quantity may be equal to an observed value, or interpolated from among two values.

Arithmetic Mean (AM) - A number equal to the average value of a population or sample. Usually obtained by summing all the values in the sample and dividing by the number of values (i.e., sample size).

Assessment Endpoint - The specific expression of the population or ecosystem that is to be protected. It can be characterized both qualitatively and quantitatively in the risk assessment.

Central Tendency Exposure (CTE) - A risk descriptor representing the average or typical individual in the population, usually considered to be the arithmetic mean or median of the risk distribution.

Credible Interval - A range of values that represent plausible bounds on a population parameter. Credible intervals may describe a parameter of an input variable (e.g., mean ingestion rate) or output variable (e.g., 95th percentile risk). The term is introduced as an alternative to the term confidence interval when the methods used to quantify uncertainty are not based entirely on statistical principles such as sampling distributions or Bayesian approaches. For example, multiple estimates of an arithmetic mean may be available from different studies reported in the literature—using professional judgment, these estimates may support a decision to describe a range of possible values for the arithmetic mean.

CTE Risk - The estimated risk corresponding to the central tendency exposure.

Cumulative Distribution Function (CDF) - Obtained by integrating the PDF or PMF, gives the cumulative probability of occurrence for a random independent variable. Each value *c* of the function is the probability that a random observation *x* will be less than or equal to *c*.

Exposure Point Concentration (EPC) - The average chemical concentration to which receptors are exposed within an exposure unit. Estimates of the EPC represent the concentration term used in exposure assessment.

Frequency Distribution/Histogram - A graphic (plot) summarizing the frequency of the values observed or measured from a population. It conveys the range of values and the count (or proportion of the sample) that was observed across that range.

High-end Risk - A risk descriptor representing the high-end, or upper tail of the risk distribution, usually considered to be equal to or greater than the 90th percentile.

Low-end Risk - A risk descriptor representing the low-end, or lower tail of the risk distribution, such as the 5th or 25th percentile.

Parameter - A value that characterizes the distribution of a random variable. Parameters commonly characterize the location, scale, shape, or bounds of the distribution. For example, a truncated normal probability distribution may be defined by four parameters: arithmetic mean [location], standard deviation [scale], and min and max [bounds]. It is important to distinguish between a variable (e.g., ingestion rate) and a parameter (e.g., arithmetic mean ingestion rate).

Probability Density Function (PDF) - A function representing the probability distribution of a continuous random variable. The density at a point refers to the probability that the variable will have a value in a narrow range about that point.

Probability Mass Function (PMF) - A function representing the probability distribution for a discrete random variable. The mass at a point refers to the probability that the variable will have a value at that point.

Reasonable Maximum Exposure (RME) - The highest exposure that is reasonably expected to occur at a site (U.S. EPA, 1989). The intent of the RME is to estimate a conservative exposure case (i.e., well above the average case) that is still within the range of possible exposures.

EXHIBIT 3-2

DEFINITIONS FOR CHAPTER 3—*Continued*

Sensitivity Analysis - Sensitivity generally refers to the variation in output of a model with respect to changes in the values of the model's input(s). Sensitivity analysis can provide a quantitative ranking of the model inputs based on their relative contributions to model output variability and uncertainty. Common metrics of sensitivity include:

- Pearson Correlation Coefficient - A statistic *r* that measures the strength and direction of linear association between the values of two quantitative variables. The square of the coefficient (r^2) is the fraction of the variance of one variable that is explained by the variance of the second variable.
- Sensitivity Ratio - Ratio of the change in model output per unit change in an input variable; also called *elasticity*.
- Spearman Rank Order Correlation Coefficient - A "distribution free" or nonparametric statistic *r* that measures the strength and direction of association between the ranks of the values (not the values themselves) of two quantitative variables. See Pearson (above) for r^2.

Target Population - The set of all receptors that are potentially at risk. Sometimes referred to as the "population of concern". A sample population is selected for statistical sampling in order to make inferences regarding the target population (see Appendix B, Section B.3.1, Concepts of Populations and Sampling).

Figure 3-1 shows a hypothetical example of an input distribution for drinking water ingestion rate. Assume that survey data for drinking water ingestion rates were compiled in order to select and fit a probability distribution. One of the first steps in exploring the data set may be to plot a frequency distribution. In the graph, the height of the bars (the y-axis) represents the relative frequency of ingestion rates in the population and the spread of the bars (the x-axis) is the varying amounts of water ingested (L/day). Since ingestion rate is a continuous random variable, the probability distribution can also be represented graphically with a probability density function (PDF). Assume that the following parameters are estimated from the sample: arithmetic mean=1.36, standard deviation=0.36, geometric mean=1.31, and geometric standard deviation=1.30. These parameter estimates may be used to define a variety of probability distributions, including a 2-parameter lognormal distribution. The fit of the lognormal distribution can be evaluated by visual inspection using the PDF given by Figure 3-1, or by a lognormal probability plot (see Appendix B).

The y-axis for a PDF is referred to as the *probability density*, where the density at a point on the x-axis represents the probability that a variable will have a value within a narrow range about the point. This type of graph shows, for example, that there is a greater area under the curve (greater probability density) in the 1-2 L/day range than 0-1 L/day or 2-3 L/day. That is, most people reported consuming 1-2 L/day of drinking water. By selecting a lognormal distribution to characterize inter-individual variability, we can state more precisely that 1 L/day corresponds to the 15th percentile and 2 L/day corresponds to the 95th percentile, so approximately 80% (i.e., 0.95–0.15=0.80) of the population is likely to consume between 1 and 2 L/day of drinking water.

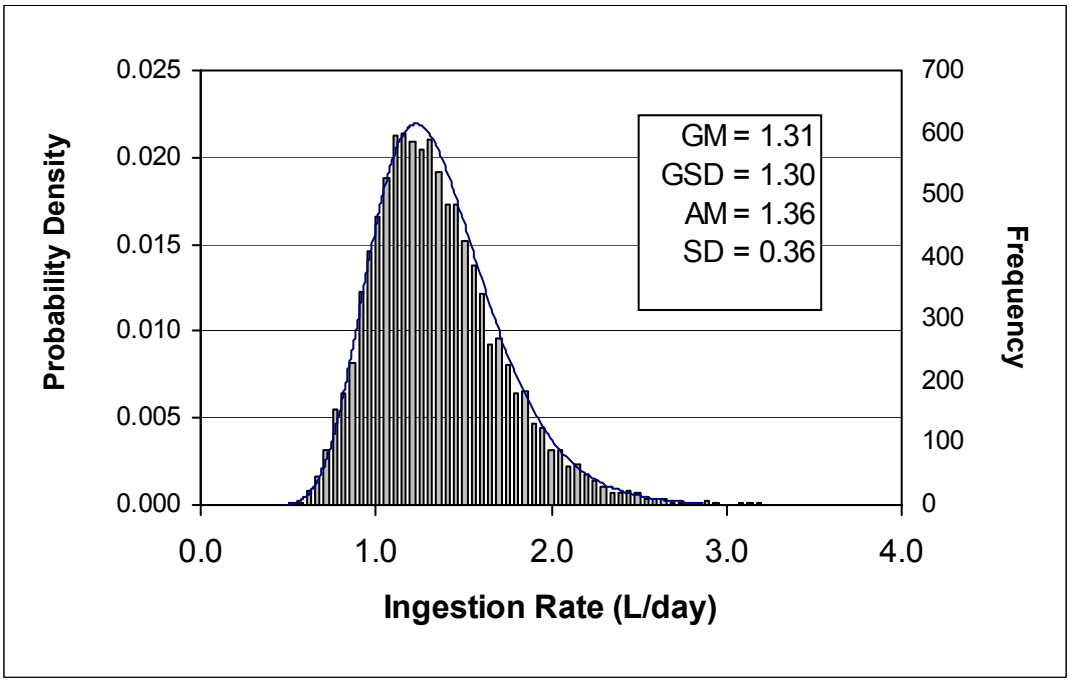


Figure 3-1. Example of a frequency distribution for adult drinking water ingestion rates, overlaid by a graph of the probability density function (PDF) for a lognormal distribution defined by the sample statistics. The distribution represents inter-individual variability in water intakes and is characterized by two parameters. Typically, the geometric mean (GM) and geometric standard deviation (GSD), or the arithmetic mean (AM) and arithmetic standard deviation (SD) are presented to characterize a lognormal distribution.

3.1.1 DEVELOPING DISTRIBUTIONS FOR EXPOSURE VARIABLES

When site-specific data or representative surrogate data are available, a probability distribution can be fit to that data to characterize variability. Appendix B describes how to fit distributions to data, how to assess the quality of the fit and discusses topics such as the sensitivity of the tails of the distribution to various PDFs, and correlations among variables. Many of the issues discussed below regarding the use of site-specific data or surrogate data are relevant to both point estimate risk assessment and PRA.

For the majority of the exposure variables, such as exposure duration, water intake rates, and body weight, site-specific data will not be available. The risk assessor will have to either select a distribution from existing sources, or develop a distribution from published data sets and data summaries. Examples of sources for these distributions and data sets are EPA's *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), Oregon Department of Environmental Quality's *Guidance for Use of Probabilistic Analysis in Human Health Risk Assessment* (Oregon DEQ, 1998), and the scientific literature. An appropriate PDF should be determined in collaboration with the regional risk assessor. The process by which PDFs are to be selected and evaluated should be described in the workplan. EPA's Superfund program is in the process of developing a ranking methodology to evaluate data representativeness relevant to various exposures scenarios. Following peer review and project completion, the results will be posted on EPA Superfund web page.

> ☞ *At this time, EPA does not recommend generic or default probability distributions for exposure variables.*

Regardless of whether a PDF is derived from site-specific measurements or obtained from the open literature, the risk assessor should carefully evaluate the applicability of the distribution to the target population at the site. The distribution selected should be derived from the target population or from a surrogate population that is representative of the target population at the site. For example, a distribution based on homegrown vegetable consumption in an urban population would not be representative for a farming population in the Midwest. If such a distribution were to be used, (and no other data were available), the uncertainty and bias that this PDF would impart to the risk estimate should be communicated to the risk decision makers.

For purposes of risk management decision making, the significance of not having site-specific data should be evaluated in the context of representativeness and sensitivity analysis. If published data are representative of the potentially exposed population, then site-specific data may be unnecessary. For example, body weights of children and adults have been well studied from national surveys and can generally be considered reasonable surrogates for use in site risk assessments. Furthermore, even if a variable is likely to vary among different exposed populations, it may not contribute greatly to the variance or uncertainty in risk estimates. In this case, surrogate data may also be used with confidence in the risk estimate. In addition, the PRA may be simplified by using point estimates instead of probability distributions for the "less sensitive" exposure variables. In part, the decision to use a point estimate in lieu of a probability distribution must balance the benefit of simplifying the analysis and the communication process (see Chapter 6), against the reduction (however small) in the variance of the risk distribution. The utility of sensitivity analysis in identifying the important factors in a risk estimate is discussed further below and in Appendix A.

It is also important to evaluate the sample design and sample size when deciding to apply a distribution to a specific site. Depending on the situation, a very large data set derived from a national

population may be more useful than a site-specific data set derived from a small, incomplete, or poorly designed study. Appendix B provides additional discussion on how to evaluate the data and studies that form the basis for a distribution. Often, the question arises regarding the appropriateness of combining data sets to derive a PDF. Before combining data sets, a careful evaluation should be made of the representativeness of the study populations, and the similarity in study designs and quality. In addition, statistical tests may be used to determine whether or not data sets are compatible with a common probability distribution (Hedges and Olkin, 1985; Stiteler et al., 1993). In general, risk assessors should be reluctant to combine data sets for the purpose of developing a PDF that characterizes variability. Due to the number of potential differences inherent in the study design, alternative data sets may provide a better measure of uncertainty in the probability distribution and parameter estimates, rather than a means of increasing the overall sample size for defining a single probability distribution. For example, if multiple data sets are available, a more informative approach may be to incorporate each data set into the PRA in a separate analysis, as a form of sensitivity analysis on the choice of alternative data sets.

Each probability distribution used in a Monte Carlo Analysis (MCA) should be presented with sufficient detail that the analysis can be reproduced (see Chapter 1, Section 1.4, Condition #2). This information may be presented in tabular and/or graphical summaries. Important information for a summary table would include a description of the distribution type (e.g., lognormal, gamma, etc.), the parameters that define the distribution (e.g., mean and standard deviation, and possibly upper and lower truncation limits for a normal distribution), units, and appropriate references (see Table 3-6, for example). The table should also indicate whether the distribution describes variability or uncertainty. The report should discuss the representativeness of the data and why a particular data set was selected if alternatives were available. Graphical summaries of the distributions may include both PDFs and cumulative distribution functions (CDFs), and should generally be used to document distributions that characterize site-specific data.

3.1.2 CHARACTERIZING RISK USING PRA

Quantitative risk characterization involves evaluating exposure (or intake) estimates against a benchmark of toxicity, such as a cancer slope factor or a noncancer hazard quotient. The general equation used for quantifying cancer risk from ingestion of contaminated soil is shown in Exhibit 3-3, and the equation for noncarcinogenic hazard is shown in Exhibit 3-4. A Hazard Index is equal to the sum of chemical-specific Hazard Quotients.

At this time, this guidance does not propose probabilistic approaches for dose-response in human health assessment and, further, *discourages undertaking such activities on a site-by-site basis*. Such activities require contaminant-specific national consensus development and national policy development (see Chapter 1, Section 1.4.1). Chapter 4 discusses methods for applying probabilistic approaches to ecological dose-response assessment.

The probabilistic calculation of risk involves random sampling from each of the exposure variable distributions. The output of this process is a distribution of risk estimates. When the calculation of risk (or any other model endpoint) is repeated many times using Monte Carlo techniques to sample the variables at random, the resulting distribution of risk estimates can be displayed in a similar fashion. The type of summary graph used to convey the results of a MCA depends on the risk management needs. For example, Chapter 1, Figure 1-3 shows how a PDF for risk might be used to compare the probabilistic estimate of the RME risk (e.g., 95th percentile) with a risk level of concern. This type of summary can also be used to effectively illustrate the relationship between the RME risk determined from point estimate and probabilistic approaches.

EXHIBIT 3-3

EQUATION FOR CANCER RISK

$$Risk = Dose \times CSF$$

Example for Soil Ingestion

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

where,

C	=	concentration in soil (mg/kg)	ED	=	exposure duration (years)
IR	=	soil ingestion rate (mg/day)	BW	=	body weight (kg)
CF	=	conversion factor (1E-06 kg/mg)	AT	=	averaging time (days)
EF	=	exposure frequency (days/year)	CSF	=	oral cancer slope factor $(mg/kg\text{-}day)^{-1}$

EXHIBIT 3-4

EQUATION FOR NONCANCER HAZARD QUOTIENT

$$Hazard\ Quotient = \frac{Dose}{RfD}\ or\ \frac{Concentration}{RfC}$$

where,

RfD	=	reference dose, oral or dermally adjusted (mg/kg-day)
RfC	=	reference concentration, inhalation ($\mu g/m^3$)

In addition, the CDF can be especially informative for illustrating the percentile corresponding to a particular risk level of concern (e.g., cancer risk of 1E-04 or Hazard Index of 1). Figure 3-2 illustrates both the PDF and CDF for risk for a hypothetical scenario. Factors to consider when applying the PDF or CDF are discussed in Chapter 1, Exhibit 1-3. When in doubt about the appropriate type of summary to use, both the PDF and CDF should be provided for all risk distributions. At a minimum, each summary output for risk should highlight the risk descriptors of concern (e.g., 50^{th}, 90^{th}, 95^{th}, and 99.9^{th} percentiles). It can also be informative to include the results of the point estimate analysis—the risks corresponding to the central tendency exposure (CTE) and the reasonable maximum exposure (RME).

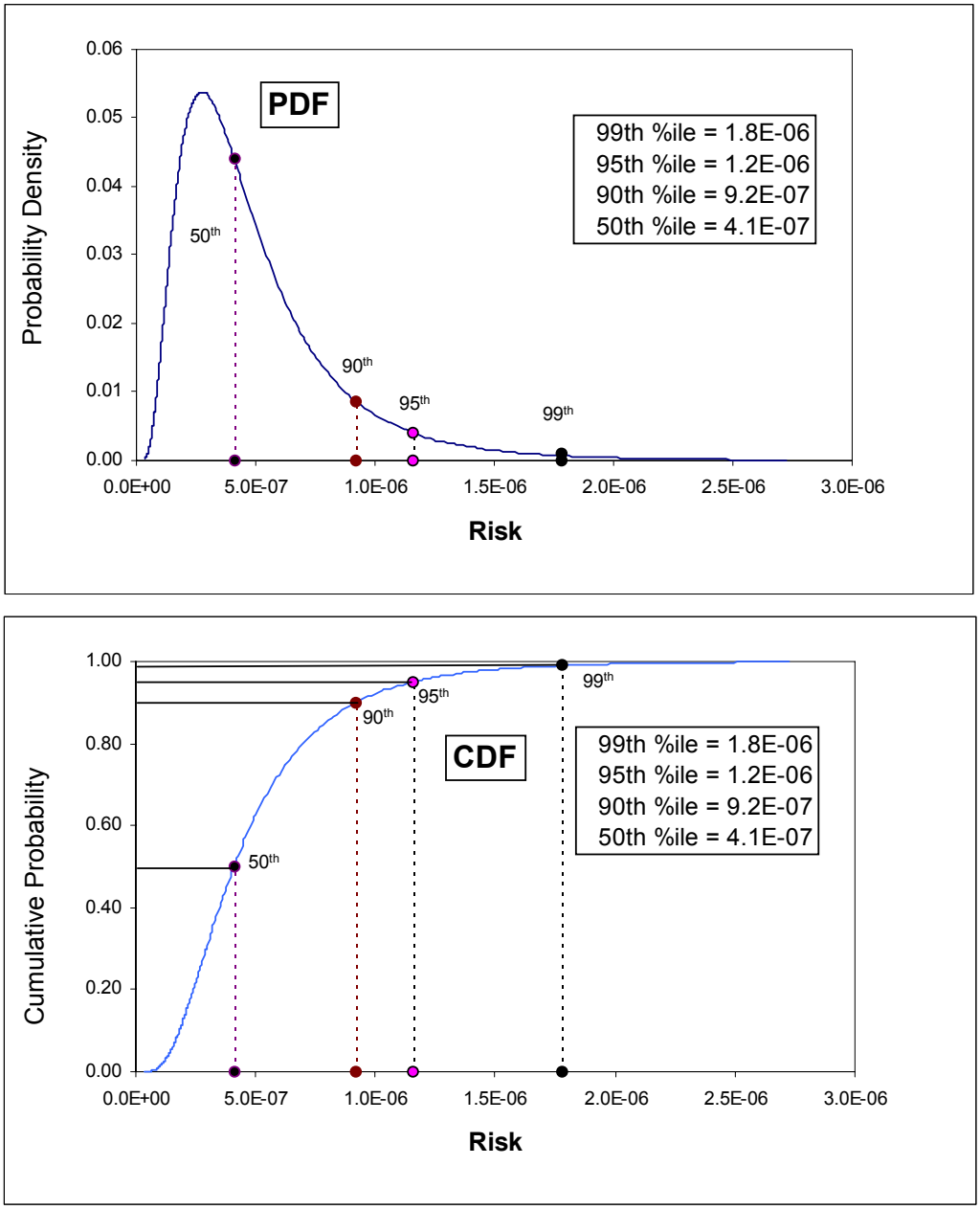


Figure 3-2. Hypothetical PRA results showing a PDF (top panel) and CDF (bottom panel) for cancer risk with selected summary statistics. The CDF rises to a maximum cumulative probability of 1.0. The CDF clearly shows that the level of regulatory concern chosen for this example (1E-06) falls between the 90th and 95th percentiles of the risk distribution.

3.2 ROLE OF THE SENSITIVITY ANALYSIS

Prior to conducting a PRA, it is worthwhile to review several points pertaining to the sensitivity analysis. As shown in Chapter 2 (Figures 2-1 and 2-2), sensitivity analysis can play an important role in decision making at each tier of the tiered process. Beginning with Tier 1, a point estimate for risk should be calculated prior to conducting a PRA. Based on the results of the point estimate, the risk assessor and risk decision makers should determine whether a probabilistic analysis will offer additional benefit. One factor in this decision may be the results of a sensitivity analysis. A primary objective of the sensitivity analysis is to determine which variables and pathways most strongly influence the risk estimate. At many Superfund sites, an estimate of cumulative risk considers contamination in multiple media, moving through multiple pathways and interacting with a number of receptors. Depending on the complexity of the site, and the modeling approaches, a risk assessment may involve one exposure pathway and few variables, or multiple pathways with many variables (e.g., multimedia fate and transport models). However, resources and time are often limited. The sensitivity analysis is invaluable in focusing these limited resources on the most influential variables and pathways.

Several methods for conducting sensitivity analysis are described in Appendix A. It is important to note that when a sensitivity analysis is performed and the major variables are identified, this does not mean that the less influential pathways and variables should be eliminated from the risk assessment. It means that because they are not major contributors to the variability or uncertainty in risk, they can be described with point estimates without affecting the risk management decision. If distributions are readily available for these less influential variables, one may use distributions. The key goal is to provide a comprehensive risk characterization that is scientifically credible and sufficient for risk decision making. The time and effort required to achieve various levels of complexity should be weighed against the value of the information provided to the risk managers.

Additionally, if a variable is specified as influential in the sensitivity analysis, this does not automatically mean that a distribution has to be developed for this variable. If the risk assessor feels that data are simply not sufficient from which to develop a distribution, then a plausible point estimate can be used. The risk assessor should be aware of a possible problem arising from using point estimates in the absence of data adequate to support a distribution. If a variable has the potential to significantly impact the risk outcome, and a very high-end or low-end point estimate is used in the PRA, this has the potential to right-shift or left-shift the final distribution of risk. Even though there might not be enough data to develop a distribution of variability for an influential variable, it would be prudent to communicate the importance of this data gap to the risk decision makers, and perhaps run multiple simulations with several plausible input distributions for that variable. Communication of this uncertainty may persuade the risk decision makers to collect additional data to better define the variable.

3.3 EXPOSURE POINT CONCENTRATION TERM

A brief discussion of the concentration term is provided below. A more complete discussion of the concentration term in PRA is provided in Appendix C. The reader is also referred to Chapter 5 on development of PRGs.

The major source of uncertainty in Superfund risk assessments is often incomplete knowledge of the concentration of one or more chemicals in various exposure media. In any risk assessment, the derivation of the concentration term will reflect assumptions about: (1) properties of the contaminant, (2) the spatial and temporal variability in contamination, (3) the behavior of the receptor, and (4) the time scale of the toxicity of the chemical(s).

Contaminant concentrations contacted by a receptor are likely to vary depending on the spatial variability of contamination and the movements of the receptor. Different individuals may be exposed to different concentrations based on inter-individual variability in activity patterns. If information regarding activity patterns is unavailable, receptors are typically assumed to exhibit random movement such that there is an equal probability of contacting any area within the exposure unit (EU). An EU is defined as the geographical area in which a receptor moves and contacts contaminated medium during the period of the exposure duration. In addition, in Superfund risk assessments, the toxicity criteria are often based on health effects associated with chronic exposure (e.g., lifetime risk of cancer following chronic daily intake over a period of 30 years). Hence, the most appropriate expression for the concentration term, for the majority of risk assessments, is one that characterizes the long-term average exposure point concentration within the EU.

> ☞ *The most appropriate expression of the exposure point concentration term for chronic exposure will characterize the long-term average concentration experienced by a receptor within the exposure unit.*

In point estimate risk assessments, the exposure point concentration term is usually calculated as the 95% upper confidence limit (95% UCL) of the arithmetic mean because of the uncertainty associated with estimating the true (i.e., population) mean concentration at a site. If the sampling density is sparse relative to the size of the EU, the uncertainty may be high due to the relatively small number of measurements available to estimate the mean concentration within the EU. The decision to use the upper confidence limit to define the concentration term introduces a measure of protectiveness by reducing the chance of underestimating the mean. Although there will be situations in which modeling variability in concentration will be the appropriate choice (e.g., non-random movement within an EU, acute exposure events, migration of groundwater contaminant plume, migration of fish, etc.), in most cases, characterization of the concentration term will focus on uncertainty. Appendix C provides a more complete discussion on characterizing both variability and uncertainty in the concentration term. Table 3-1 summarizes a number of appropriate methods for characterizing uncertainty in the parameter of an exposure variable, such as the arithmetic mean of the concentration term.

3.4 CHARACTERIZING UNCERTAINTY IN EXPOSURE VARIABLES

Uncertainty is described as a lack of knowledge about factors affecting exposure or risk. To evaluate regulatory options, risk assessors are expected to translate the available evidence, however tentative, into a probability of occurrence of an adverse health effect. Data from a sample or surrogate population are used to develop estimates of exposure and risk in a specific target population (see Section 3.1.4 and Appendix B, Section B.3.1). This extrapolation requires assumptions and inferences that have inherent strengths and limitations, and may bias the outcome of the risk estimate. For example, a common assumption in risk assessments for carcinogens is that a contaminant concentration within the boundaries of a hazardous waste site represents the concentration that a receptor is exposed to throughout the period of exposure, with the corresponding dose averaged over the course of a lifetime. This assumption may be conservative (i.e., result in overestimation of exposure) if it is unlikely that receptors will be exposed at the hazardous waste site for the entire exposure duration. It is incumbent on the risk assessor to clearly present the rationale for the assumptions used in a risk assessment, as well as their implications and limitations.

U.S. EPA guidance, including the *Exposure Assessment Guidelines* (U.S. EPA, 1992a), *Exposure Factors Handbook* (U.S. EPA, 1997a,b,c), and *Guiding Principles for Monte Carlo Analysis* (U.S. EPA, 1997d) have classified uncertainty in exposure assessment into three broad categories:

(1) *Parameter uncertainty* - uncertainty in values used to estimate variables of a model;
(2) *Model uncertainty* - uncertainty about a model structure (e.g., exposure equation) or intended use; and
(3) *Scenario uncertainty* - uncertainty regarding missing or incomplete information to fully define exposure.

Each source of uncertainty is described in detail below, along with strategies for addressing them in PRA.

3.4.1 PARAMETER UNCERTAINTY

Parameter uncertainty may be the most readily recognized source of uncertainty that is quantified in site-specific risk assessments at hazardous waste sites. Parameter uncertainty can occur in each step of the risk assessment process from data collection and evaluation, to the assessment of exposure and toxicity. Sources of parameter uncertainty may include systematic errors or bias in the data collection process, imprecision in the analytical measurements, and extrapolation from surrogate measures to represent the parameter of interest. For example, soil data collected only from the areas of highest contamination, rather than the entire area that a receptor is expected to come into contact, will result in a biased estimate of exposure.

In general, parameter uncertainty can be quantified at any stage of the tiered process, including point estimate analysis (Tier 1), one-dimensional Monte Carlo analysis (1-D MCA) (Tier 2), and two-dimensional Monte Carlo analysis (2-D MCA) (Tier 3). In the point estimate approach, parameter uncertainty may be addressed in a qualitative manner for most variables. For example, the uncertainty section of a point estimate risk assessment document might state that an absorption fraction of 100% was used to represent the amount of contaminant in soil absorbed from the gastrointestinal (GI) tract, and as a result, the risk estimate may overestimate actual risk. In addition, a sensitivity analysis may be performed, wherein one input variable at a time is changed, while leaving the others constant, to examine the effect on the outcome. In the case of absorption from the GI tract, different plausible estimates of the

high-end, or RME absorption fraction might be used as inputs to the risk equation. The differences in the risk estimates would reflect uncertainty in the RME absorption fraction.

Quantitative approaches for characterizing parameter uncertainty in exposure variables in a Monte Carlo simulation are summarized in Table 3-1. If uncertainty in only a few parameter values is of interest, multiple 1-D MCA simulations can yield the same results as a 2-D MCA simulation, but without the time and effort of a 2-D MCA. An example illustrating this concept is given in Table 3-2. With multiple 1-D MCA simulations, variability is characterized in one or more variables using probability distributions for variability (PDFv's), and uncertainty in a parameter is characterized with a series of different point estimates from a probability distribution for uncertainty (PDFu) (e.g., 95% lower confidence limit LCL [95% LCL], sample mean, and 95% UCL). In a 2-D MCA simulation, variability is characterized in one or more variables using PDFv's, and uncertainty in one or more parameters is characterized with PDFu's. With both approaches, the influence of the parameter uncertainty can be presented as a credible interval or confidence interval (CI) around the risk distribution, depending on how the PDFu's are defined. When only a few sources of parameter uncertainty are quantified, multiple 1-D MCA simulations are preferred over a 2-D MCA because the approach is easier to use and communicate. However, if the goal is to explore the effect that many sources of parameter uncertainty may have on the risk estimates simultaneously, a 2-D MCA is preferred. Iterative 1-D MCA simulations with different combinations of confidence limits may be impractical.

Table 3-1. Methods for Characterizing Parameter Uncertainty with Monte Carlo Simulations.

Approach	Example of Model Input	Method	Example of Model Output
Single Point Estimate	• 95% UCL	1-D MCA	PDFv[1] for risk, calculated using the 95% UCL for one parameter.
Multiple Point Estimates	• 95% LCL • sample mean • 95% UCL	1-D MCA	Three PDFv's for risk, representing the 90% CI for each percentile of the risk distribution.[2] The 90% CI only accounts for uncertainty in a single parameter (not multiple parameters).
Parametric PDFu[1]	PDFu for the mean based on the sampling distribution, derived from a Student's *t*-distribution.	2-D MCA	One PDFv for risk with confidence intervals at each percentile of the risk distribution. The CI reflects uncertainty in one or more parameters.
Non-parametric PDFu	PDFu for the mean based on bootstrap resampling methods.	2-D MCA	Same as parametric probability distribution for uncertainty.

[1]Probability distribution for uncertainty (PDFu) and probability distribution for variability (PDFv).
[2]The 95% UCL for the concentration term represents a 1-sided confidence interval (CI), meaning there is a 95% probability that the value is *greater* than or equal to the mean. Similarly, the 95% LCL would represent the 1-sided CI in which there is a 95% probability that the value is *less* than or equal to the mean. Both values are percentiles on the probability distribution for uncertainty (PDFu), also called the sampling distribution for the mean. Together, the 95% LCL and 95% UCL are equal to the 2-sided 90% confidence interval only for cases in which the PDFu is symmetric. For example, the sampling distribution for the arithmetic mean of a sample from a normal distribution with an unknown variance is described with the symmetric Student's *t*-distribution, whereas the PDFu for the mean of a lognormal distribution is asymmetric. In order to compare the results of multiple 1-D MCA simulations and a 2-D MCA simulation, the same methodology should be employed to define the PDFu and the corresponding confidence limits.

It is generally incorrect to combine a PDFu for one parameter (e.g., mean of the concentration term) with one or more PDFv's in other exposure factors when conducting a 1-D MCA for variability.

However, distributions for uncertainty and variability may be appropriately combined in a 2-D MCA. As discussed in Appendix D, with 2-D MCA, a clear distinction should be made between probability distributions that characterize variability (PDFv) and parameter uncertainty (PDFu). A 2-D MCA propagates the uncertainty and variability distributions separately through an exposure model, thereby making it possible to evaluate the effect of each on the risk estimates.

Example: Comparison of Multiple Point Estimates of Uncertainty in 1-D MCA, and Distributions of Uncertainty in 2-D MCA

Table 3-2 illustrates an application of the approaches presented in Table 3-1 for quantifying variability and parameter uncertainty. This is a hypothetical example, and no attempt was made to use standard default assumptions for exposure variables. Two sources of variability are quantified: (1) inter-individual variability in exposure frequency (EF), characterized by a triangular distribution, and (2) inter-individual variability in exposure duration (ED), characterized by a truncated lognormal distribution. In addition, two sources of uncertainty are presented: (1) a point estimate for soil and dust ingestion rate, intended to characterize the RME; and (2) an upper truncation limit of the lognormal distribution for ED, intended to represent a plausible upper bound for the exposed population. Methods for quantifying these sources of uncertainty are discussed below. Additional sources of uncertainty may also have been explored. For example, the choice of a triangular distribution for a PDFv may be provocative for some risk assessors, since there are few cases in which empirical data suggest a random sample is from a triangular distribution. Nevertheless, triangular distributions may be considered rough, or "preliminary" distributions (see Chapter 2 and Appendix B, Section B.2) for cases when the available information supports a plausible range and central tendency.

The choice of distributions is a potential source of uncertainty that can be explored by rerunning simulations with each alternative, plausible choice, and examining the effect on the RME risk. Simulations with preliminary simulations may yield at least three different outcomes. First, this type of sensitivity analysis can help guide efforts to improve characterizations of variability for selected variables that have the greatest affect on the risk estimates. Second, results may provide justification to exit the tiered process without continuing with additional Monte Carlo simulations since further effort would be unlikely to change the risk management decision. Finally, if the major sources of uncertainty can be clearly identified, a subset of the less sensitive variables may be defined by point estimates without significantly reducing the uncertainty in the risk estimates.

Parameter uncertainty can be quantified for both point estimates and PDFv's. In this example, both types of inputs (i.e., point estimates and PDFv's) are presented as sources of parameter uncertainty: the RME point estimate for soil and dust ingestion rate (IRsd), and the upper truncation limit on a PDFv for ED. For IRsd, assume that three different studies provide equally plausible values for the RME: 50, 100, and 200 mg/day. A uniform PDFu is specified to characterize this range of plausible values. For ED, assume that the maximum value reported from a site-specific survey was 26 years, but surrogate data for other populations suggest the maximum may be as long as 40 years. A uniform PDFu is specified to characterize this range of plausible values as well.

In Cases 1-3, the impact of uncertainty in IRsd and ED was evaluated using a series 1-D MCA simulations. Inputs for uncertain parameters associated with IRsd and ED in Case 1, 2, and 3 represent the minimum, central tendency, and maximum values, respectively. Each simulation yields a different risk distribution based on different combinations of point estimates for parameters. Although a PDFu was specified for IRsd, it would have been incorrect to combine the PDFu with the PDFv's for EF and ED in a

1-D MCA because the result would have been a single distribution of risk that co-mingled uncertainty and variability.

In Case 4, a single 2-D MCA simulation was run using the PDFu's for uncertainty and the PDFv's for variability. By propagating variability and uncertainty separately, the 2-D MCA yields a series of distributions of risk, from which credible intervals can be calculated for each percentile of the CDF.

$$Risk = \frac{C \times IR \times CF \times EF \times ED}{BW \times AT} \times CSF_{oral}$$

Table 3-2. Example of 1-D MCA and 2-D MCA.

Variable	Type of Input	1-D MCA			2-D MCA
		Case 1	Case 2	Case 3	Case 4
C (mg/kg)	pt estimate	500	500	500	500
IRsd (mg/day)	pt estimate	50	100	200	see below
	PDFu for pt estimate	--	--	--	uniform (50, 200)[a]
CF (kg/mg)	pt estimate	1E-06	1E-06	1E-06	1E-06
EF (days/year)	PDFv	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350	triangular min = 200 mode = 250 max = 350
ED (years)	PDFv	T-lognormal mean = 9 stdv = 10 max = 26	T-lognormal mean = 9 stdv = 10 max = 33	T-lognormal mean = 9 stdv = 10 max = 40	T-lognormal mean = 9 stdev = 10 max = PDFu (see below)
	PDFu for parameter of PDFv	--	--	--	max ~ uniform (26, 40)[b]
BW (kg)	pt estimate	70	70	70	70
AT (days)	pt estimate	25550	25550	25550	25550
CSF $(mg/kg\text{-}day)^{-1}$	pt estimate	1E-01	1E-01	1E-01	1E-01

[a]Uncertainty in the RME point estimate, defined by a uniform distribution with parameters (minimum, maximum).
[b]Uncertainty in the upper truncation limit of the lognormal distribution, defined by a PDFv with parameters (mean, standard deviation, maximum) and a PDFu for the maximum defined by a uniform distribution with parameters (minimum, maximum).

Monte Carlo Simulation Results

Figures 3-3 and 3-4 illustrate CDFs for risk produced from Monte Carlo simulations using *Crystal Ball*® 2000. The 1-D MCA simulations (Figure 3-3) were run with 10,000 iterations and Latin Hypercube sampling. The 2-D MCA simulation (Figure 3-4) was run with 250 iterations of the outer loop (uncertainty) and 2,000 iterations of the inner loop (variability). Details regarding 2-D MCA simulation are given in Appendix D.

Figure 3-3 shows CDFs for risk based on three simulations of a 1-D MCA simulation. Each simulation used a different combination of plausible estimates of the RME value for IRsd and the upper truncation limit for ED, as discussed above. The results provide a bounding estimate on the risk distribution given these two sources of uncertainty. The 95th percentile risk, highlighted as an example of the RME risk estimate, may range from approximately 7E-06 to 3.5E-05.

Figure 3-4 shows a single CDF for risk, representing the central tendency risk distribution. This CDF was derived by simulating uncertainty in the risk distribution using 2-D MCA. For this example, the 2-D MCA yields 250 simulations of the risk distributions for variability, so that there are 250 plausible estimates of each percentile of the risk distribution. In practice, more than 250 simulations may be needed to adequately quantify uncertainty in the risk distribution. Results of a 2-D MCA can be presented as probability distributions of uncertainty, or box-and-whisker plots of uncertainty at selected percentiles of the risk distributions. Figure 3-4 shows the central tendency (50th percentile) estimate of uncertainty for the entire CDF of risk. In addition, a box-and-whisker plot is shown at the 95th percentile of the CDF. Selected statistics for the box-and-whisker plot are included in a text box on the graphic (i.e., minimum; 5th, 50th, and 95th percentiles, and maximum). The 90% credible interval is given by the 5th and 95th percentiles. For this example, the 90% credible interval for the 95th percentile of the risk distribution is: [9.1E-06, 3.1E-05].

Figures 3-3 and 3-4 demonstrate that the two approaches (i.e., multiple 1-D MCA and 2-D MCA) can yield the same results. However, when there are numerous sources of uncertainty, 2-D MCA offers at least two advantages over multiple 1-D MCA simulations: (1) 2-D MCA allows the multiple sources of uncertainty to be included simultaneously so the approach is more efficient than a series of 1-D MCA simulations; and (2) multiple 1-D MCA simulations yield multiple estimates of the RME risk, but it is not possible to characterize the uncertainty in the RME risk in quantitative terms; a 2-D MCA yields a PDFu for RME risk, which allows for statements regarding the level of certainty that the RME risk is above or below a risk level of concern.

The 95th percentile is a focus of this example because it is a recommended starting point for determining the risk corresponding to the RME. Chapter 7 provides guidance to the risk decision makers on choosing an appropriate percentile (on a distribution of variability) within the RME risk range (90th to 99.9th percentiles). The chapter also includes a qualitative consideration of the uncertainty or confidence surrounding a risk estimate in the decision-making process.

Figure 3-3

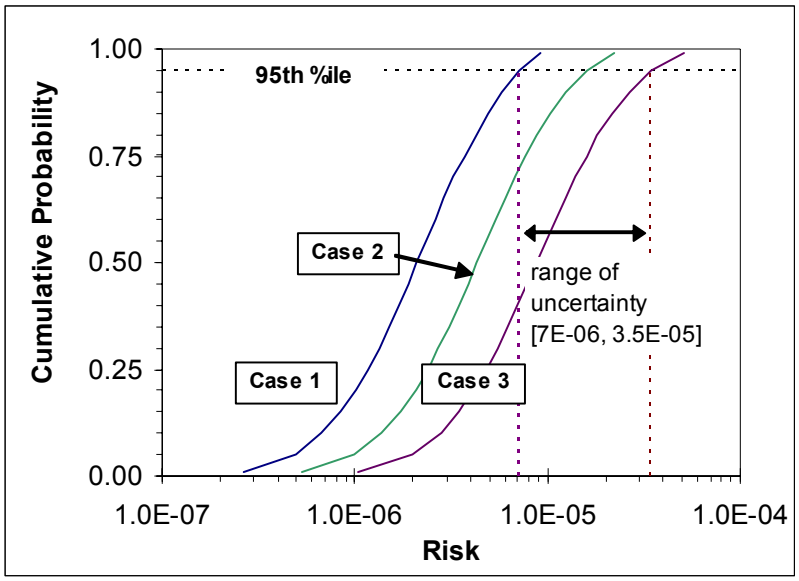
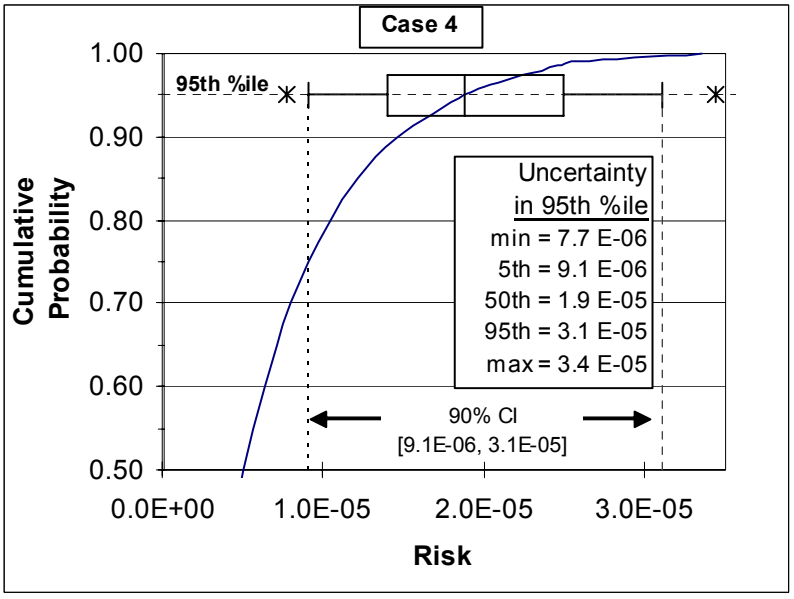


Figure 3-4



3.4.2 SCENARIO AND MODEL UNCERTAINTY

All models are simplified representations of complex biological and physical processes. As such, they, and the scenarios to which they are applied, may introduce a significant source of uncertainty into an exposure and risk estimate. Models may exclude important variables or important pathways of exposure, ignore interactions between inputs, use surrogate variables that are different from the target variables, or they may be designed for specific scenarios and not others. As a result, a model may not adequately represent all aspects of the phenomena it was intended to approximate or it may not be appropriate to predict outcomes for a different type of scenario. For example, a model intended to estimate risk from continuous, steady state exposures to a contaminant may not be appropriate or applicable for estimating risk from acute or subchronic exposure events. In any risk assessment, it is important to understand the original intent of a model, the assumptions being made in a model, what the parameters represent, and how they interact. Based on this knowledge, one can begin to understand how representative and applicable (or inapplicable) a model may be to a given scenario. If multiple models exist that can be applied to a given scenario, it may be useful to compare and contrast results in order to understand the potential implications of the differences. The use of multiple models, or models with varying levels of sophistication, may provide valuable information on the uncertainty introduced into a risk estimate as the result of model or scenario uncertainty. The collection of measured data as a reality check against a given parameter or the predicted model outcome (such as the collection of vegetable and fruit contaminant data to compare against modeled uptake into plants) is also useful in attempting to reduce or at least gain a better understanding of model and scenario uncertainty.

3.5 EXAMPLE OF PRA FOR HUMAN HEALTH

The following hypothetical example provides a conceptual walk-through of the tiered approach for PRA in Superfund risk assessment. The example begins with a baseline human health point estimate risk assessment (Tier 1) and moves to Tier 2, in which multiple iterations of a 1-D MCA are run using default and site-specific assumptions for input distributions. The general concepts associated with the tiered approach are discussed in Chapter 2, and a similar example for ecological risk assessment is given in Chapter 4. The 1-D MCA results are based on simulations with *Crystal Ball*® 2000 using 10,000 iterations and Latin Hypercube sampling. These settings were sufficient to obtain stability (i.e., <1% difference) in the 95% percentile risk estimate. The example is presented in Exhibit 3-5. Tables and figures supporting the example are given immediately following the exhibit.

EXHIBIT 3-5
USING THE TIERED PROCESS FOR PRA
HYPOTHETICAL CASE STUDY FOR HUMAN HEALTH RISK ASSESSMENT

RI Planning/Scoping/Problem Formulation/Data Collection

- Site Description: Former federal facility
- Site Size: 100 acres (5 acres within spill area (ISA); 95 acres outside spill area (OSA))
- Stakeholders: Refuge employees, environmental activists, etc.
- Land Use: Future wildlife refuge
- Receptors: Future wildlife refuge workers (i.e., ornithologists and fishery biologists)
- Sampling Data: n=35 surface soil samples (see Figure 3-5 for sample locations)
- Chemical of Concern: ChemX
- Chemical Properties: Nonvolatile
- Toxicological Properties: Carcinogen: CSF_{oral} and CSF_{dermal}= 5.5E-02, CSF_{inh}= 2.73E-02; Noncarcinogenic health data are lacking
- Risk Level of Concern: 1E-04 for cancer

↓

Tier 1 Point Estimate - Baseline Risk Assessment

- Exposure Unit: (see Figure 3-5) ornithologist (exposed in OSA) and fishery biologist (exposed in ISA)
- Exposure Pathways: Ingestion of soil/dust; inhalation of fugitive dust, dermal absorption
- Concentration Term: 95% UCL for arithmetic mean (Table 3-3)
- Risk Equations: Exhibit 3-6
- Exposure Parameters: Table 3-4
- Results: Table 3-5

↓

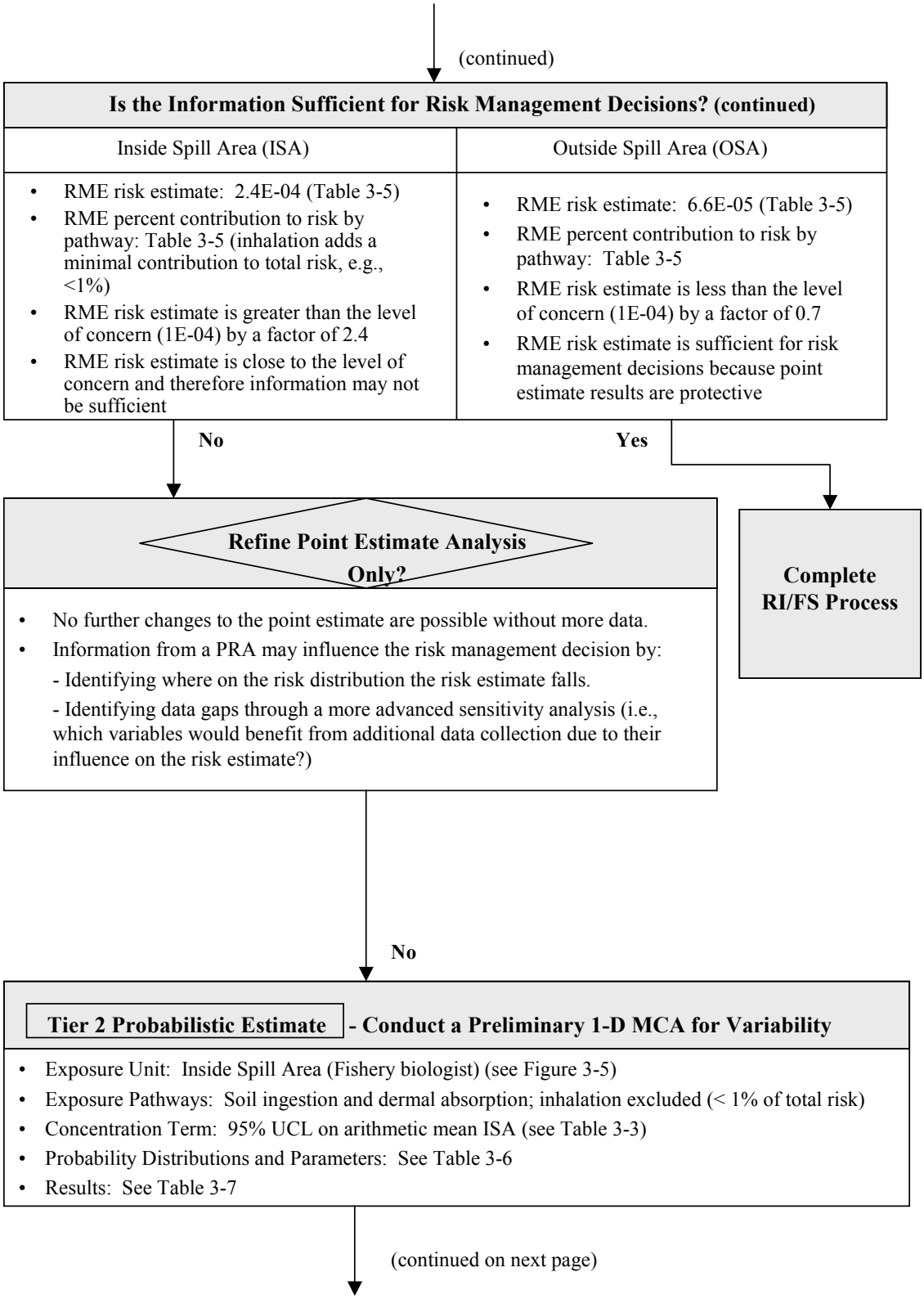
SMDP Is the Information Sufficient for Risk Management Decisions?

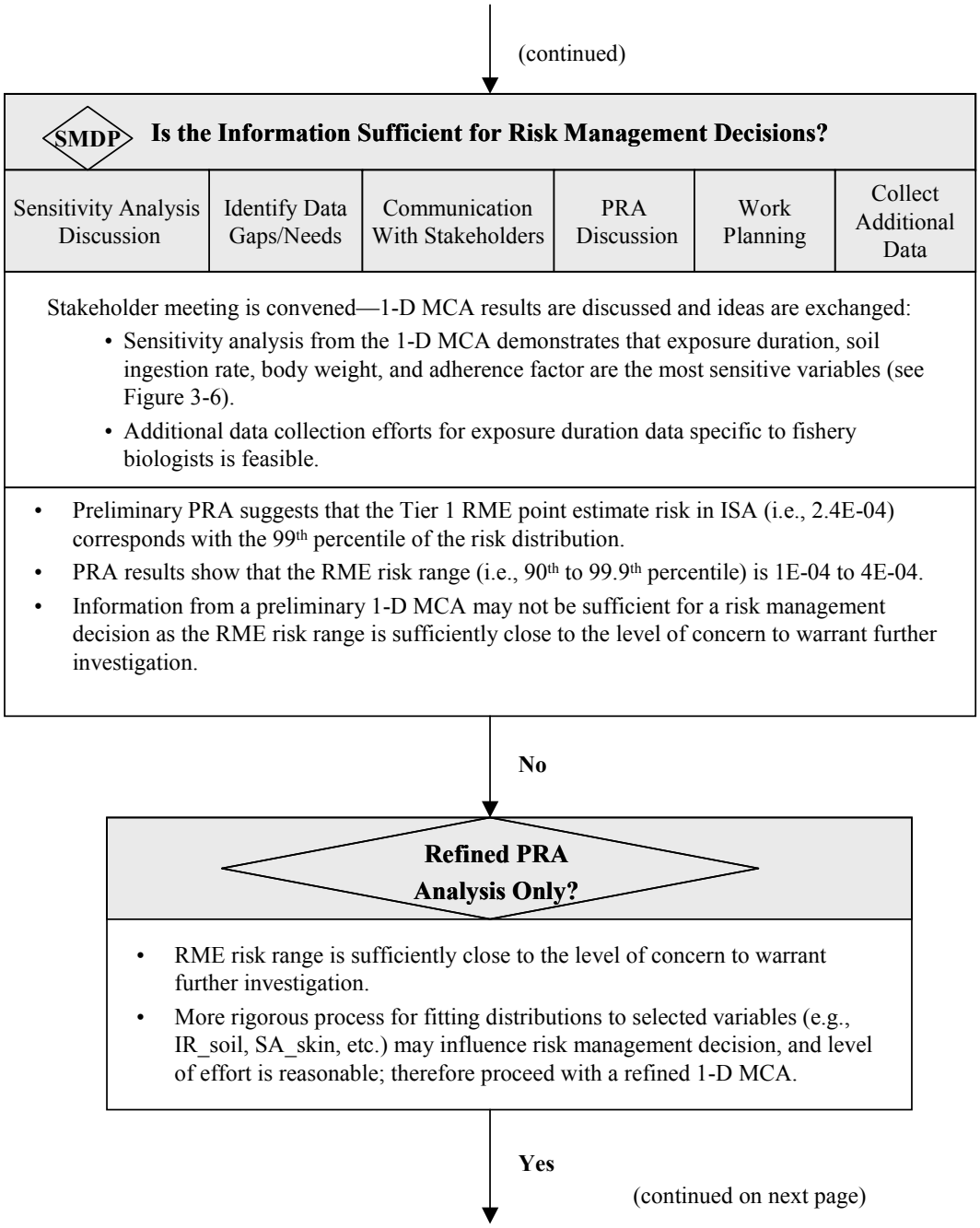
Sensitivity Analysis Discussion	Identify Data Gaps/Needs	Communication With Stakeholders	PRA Discussion	Work Planning	Collect Additional Data

Stakeholder meeting is convened—point estimate results are discussed and ideas are exchanged as follows:

- Risk estimates are expected to be conservative due to the use of standard default exposure parameters, but are the defaults representative?
- Stakeholders are concerned about risk to workers and about the consequences of remediation (e.g., negative impacts on habitat and potential job losses).
- Stakeholders are concerned about the relevance of some nonsite-specific exposure variables (e.g., exposure duration), but are not sure which variables to investigate further (i.e., which is the most influential?).
- Results of the sensitivity analysis from point estimate risk assessment cannot identify where the high end risk estimate falls on the risk distribution.
- There is sufficient information (e.g., arithmetic mean, standard deviation, percentiles) for some of the exposure variables to develop initial probability distributions to characterize variability.

(continued on next page)


(continued)
Is the Information Sufficient for Risk Management Decisions? (continued)
Inside Spill Area (ISA)
Outside Spill Area (OSA)
• RME risk estimate: 2.4E-04 (Table 3-5)
• RME percent contribution to risk by pathway: Table 3-5 (inhalation adds a minimal contribution to total risk, e.g., <1%)
• RME risk estimate is greater than the level of concern (1E-04) by a factor of 2.4
• RME risk estimate is close to the level of concern and therefore information may not be sufficient
• RME risk estimate: 6.6E-05 (Table 3-5)
• RME percent contribution to risk by pathway: Table 3-5
• RME risk estimate is less than the level of concern (1E-04) by a factor of 0.7
• RME risk estimate is sufficient for risk management decisions because point estimate results are protective
No
Yes
Refine Point Estimate Analysis Only?
Complete RI/FS Process
• No further changes to the point estimate are possible without more data.
• Information from a PRA may influence the risk management decision by:
- Identifying where on the risk distribution the risk estimate falls.
- Identifying data gaps through a more advanced sensitivity analysis (i.e., which variables would benefit from additional data collection due to their influence on the risk estimate?)
No
Tier 2 Probabilistic Estimate - Conduct a Preliminary 1-D MCA for Variability
• Exposure Unit: Inside Spill Area (Fishery biologist) (see Figure 3-5)
• Exposure Pathways: Soil ingestion and dermal absorption; inhalation excluded (< 1% of total risk)
• Concentration Term: 95% UCL on arithmetic mean ISA (see Table 3-3)
• Probability Distributions and Parameters: See Table 3-6
• Results: See Table 3-7
(continued on next page)


(continued)
SMDP
Is the Information Sufficient for Risk Management Decisions?
Sensitivity Analysis Discussion
Identify Data Gaps/Needs
Communication With Stakeholders
PRA Discussion
Work Planning
Collect Additional Data
Stakeholder meeting is convened—1-D MCA results are discussed and ideas are exchanged:
• Sensitivity analysis from the 1-D MCA demonstrates that exposure duration, soil ingestion rate, body weight, and adherence factor are the most sensitive variables (see Figure 3-6).
• Additional data collection efforts for exposure duration data specific to fishery biologists is feasible.
• Preliminary PRA suggests that the Tier 1 RME point estimate risk in ISA (i.e., 2.4E-04) corresponds with the 99th percentile of the risk distribution.
• PRA results show that the RME risk range (i.e., 90th to 99.9th percentile) is 1E-04 to 4E-04.
• Information from a preliminary 1-D MCA may not be sufficient for a risk management decision as the RME risk range is sufficiently close to the level of concern to warrant further investigation.
No
Refined PRA Analysis Only?
• RME risk range is sufficiently close to the level of concern to warrant further investigation.
• More rigorous process for fitting distributions to selected variables (e.g., IR_soil, SA_skin, etc.) may influence risk management decision, and level of effort is reasonable; therefore proceed with a refined 1-D MCA.
Yes
(continued on next page)

Yes

(continued)

Tier 2 Refined PRA - Conduct Refined 1-D MCA and Refined Point Estimate

- Exposure Unit: Fishery biologist-inside spill area (ISA) (see Figure 3-5)
- Exposure Pathways: Ingestion of soil and dust, and dermal absorption
- Concentration Term: 95% UCL on arithmetic mean
- Probability Distributions/Parameters: see Table 3-8 for sample data and summary statistics; exposure duration defined by lognormal PDF (arithmetic mean=14, SD=9.4, upper truncation of 44 years)
- Results: see Table 3-9

SMDP Is the Information Sufficient for Risk Management Decisions?

Sensitivity Analysis Discussion	Identify Data Gaps/Needs	Communication With Stakeholders	PRA Discussion	Work Planning	Collect Additional Data

Stakeholders meeting is convened. Refined 1-D MCA results are discussed and ideas are exchanged as follows:

- Sensitivity analysis from refined 1-D MCA indicates that the use of site-specific data did not significantly alter the relative ranking or magnitude of rank correlations for input variables (similar graphic as Figure 3-6).
- Refined 1-D MCA results suggest that the refined RME point estimate risk corresponds with the 99th percentile of the risk distribution (Table 3-9).
- Refined 1-D MCA results show that the RME range (i.e., 90th to 99.9th percentile) is 1.6E-04 to 5E-04, with 95th percentile of 2.1E-04.
- Information from refined 1-D MCA is sufficient for risk management decision because the RME risk (95th percentile) is above the level of concern of 1E-04 using site specific exposure duration data, and additional data collection on IR_soil term is not warranted. Complete RI/FS process.

Yes

Complete RI/FS Process

- Stakeholders and RPM decide that the best remedial alternative is to remove surface soil in the 5 acre spill area and cover the refuge area with clean fill before beginning refuge construction.

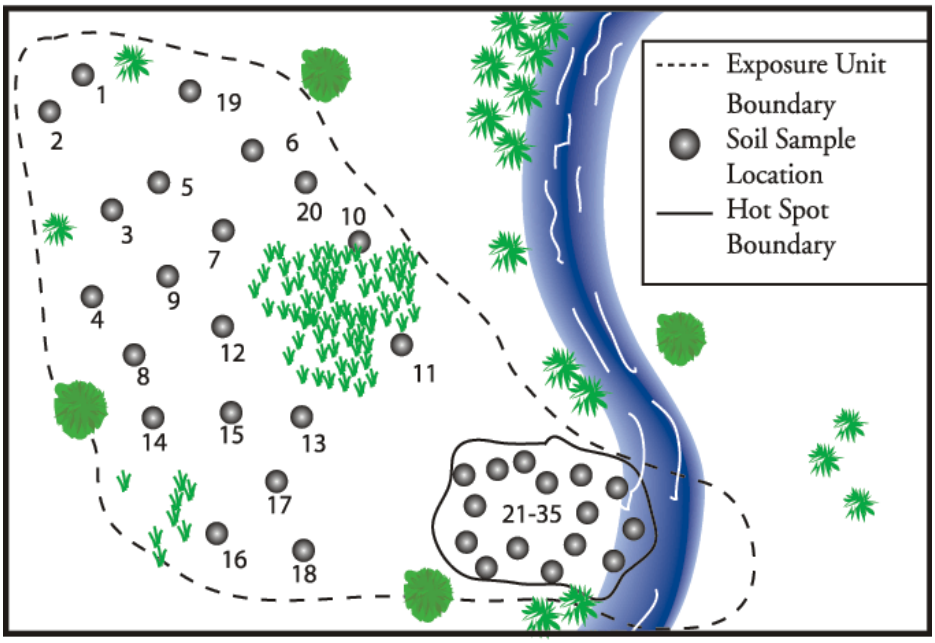


Figure 3-5. Site map for future wildlife refuge showing boundaries for the exposure unit and potential hotspot, as well as sampling locations (n=35). Sample numbers correspond with concentration data given in Table 3-3.

[1]The 95% UCL was estimated using the Land method (see Appendix C).

Table 3-3. Concentrations in Surface Soil (mg/kg).

Outside Spill Area (n=20)		Inside Spill Area (n=15)	
1088	305	1934	970
646	2787	402	985
3943	760	4215	743
149	149	1121	158
3704	1088	629	21296
845	837	2293	
488	1295	257	
387	1239	288	
1438	1006	57	
2502	283	228	

Summary Statistics	Outside Spill Area	Inside Spill Area
Mean	1247	2372
Standard Deviation	1121	5348
95% UCL[1]	2303	8444

EXHIBIT 3-6

RISK EQUATIONS

Soil Ingestion

$$\text{Risk} = \frac{Cs \times CF \times IRs \times FI \times EF \times ED}{BW \times AT} \times \text{Oral CSF}$$

Dermal Absorption

$$\text{Risk} = \frac{Cs \times CF \times SA \times AF \times ABS \times EF \times ED}{BW \times AT} \times \text{Dermal-Adjusted CSF}$$

Inhalation of Fugitive Dust

$$\text{Risk} = \frac{Cs \times 1/PEF \times IRa \times ET \times EF \times ED}{BW \times AT} \times \text{Inhalation CSF}$$

Total Risk = Sum of risks from each exposure pathway (soil + dermal + inhalation)

Where:

Cs = Concentration of ChemX in soil (mg/kg)
IRs = Soil ingestion rate for receptor (mg/day)
FI = Fraction ingested from contaminated source (unitless)
CF = Conversion factor (1E-06 kg/mg)
SA = Skin surface area available for exposure (cm^2/event)
AF = Soil to skin adherence factor for ChemX (mg/cm^2)
ABS = Absorption factor for ChemX (unitless)
IRa = Inhalation rate for receptor (m^3/hr)
PEF = Soil-to-air particulate emission factor (kg/m^3)
ET = Exposure time for receptor (hours/day)
EF = Exposure frequency for receptor (days/year)
ED = Exposure duration for receptor (years)
BW = Body weight of receptor (kg)
AT = Averaging time (years)
CSF = Cancer slope factor (oral, dermal, inhalation) $(mg/kg\text{-}day)^{-1}$

Table 3-4. Exposure Parameters used in Point Estimate Analysis.

Exposure Variable	CTE Value	RME Value	Units	Reference
IRs	50	100	mg/day	CTE: U.S. EPA, 1997a, p. 4–25 RME: U.S. EPA, 2001
FI	0.5	1	unitless	Site-specific
CF	1E-06	1E-06	kg/mg	Constant
SA	3300	3300	cm^2/event	U.S. EPA, 2001, 50th percentile value for all adult workers—exposure to face, forearms, and hands
AF	0.1	0.2	mg/cm^2	CTE: U.S. EPA, 1998; Table 3.3, value for gardeners RME: U.S. EPA, 2001
ABS	0.1	0.1	unitless	U.S. EPA, 1998, default for semi-volatile organic compounds (SVOCs)
IRa	1.3	3.3	m^3/hr	U.S. EPA, 1997a, p. 5–24, outdoor worker hourly average: mean and upper percentile
PEF	1.36E+09	1.36E+09	kg/m^3	U.S. EPA, 2001
ET	8	8	hours/day	Site-specific
EF	200	225	days/year	CTE: Site-specific assumption RME: U.S. EPA, 2001
ED	5	25	years	CTE: U.S. EPA, 1993, p. 6 RME: U.S. EPA, 2001
BW	70	70	kg	U.S. EPA, 1993, p. 7
AT	25550	25550	days	constant

CTE = central tendency exposure; RME = reasonable maximum exposure.

Table 3-5. Point Estimate Risks and Exposure Pathway Contributions**.**

Risk Estimate by Exposure Pathway	Inside Spill Area (n = 15)		Outside Spill Area (n = 20)	
	CTE	RME	CTE	RME
Soil Ingestion	6.5E-06 (43 %)	1.5E-04 (60 %)	1.7E-06 (43 %)	4.0E-05 (60 %)
Dermal Absorption	8.6E-06 (57 %)	9.6E-05 (40 %)	2.3E-06 (57 %)	2.6E-05 (40 %)
Inhalation	9.9E-10 (< 1 %)	1.4E-08 (< 1 %)	2.7E-10 (< 1 %)	3.8E-09 (< 1 %)
Total Risk	**1.5E-05**	**2.4E-04**	**4.1E-06**	**6.6E-05**

Example of % contribution: % Soil for RME risk inside spill area = (Soil risk / Total risk) x 100%
= (1.46E-04 / 2.42E-04) x 100% = 60%

Table 3-6. Input Distributions for Exposure Variables used in 1-D MCA for Variability.

Exposure Variable[1]	Distribution Type	Parameters[2]	Units	Reference
IR_soil	Triangular	0, 50, 100	mg/day	U.S. EPA, 1993, 2001
SA_skin[3]	Lognormal	18150, 37.4	cm^2	U.S. EPA, 1997a, Table 6-4 (Total male/female body surface area)
Absorption Fraction	Uniform	0.1, 0.2	mg/cm^2	U.S. EPA, 2001; minimum truncation limit is professional judgment
IR_air	Lognormal	1.68, 0.72	m^3/hour	U.S. EPA, 1996, p.5–10
EF	Triangular	200, 225, 250	days	U.S. EPA, 2001; truncation limits are professional judgment
ED	Lognormal[4]	11.7, 7.0	years	U.S. EPA, 1997b, Table 15-161 and U.S. EPA, 2001 (Mean value is based on average of total median tenure for professional specialty and farming, forestry, and fishing)
	Truncated Lognormal[5]	14.0, 9.4, 44.0	years	Site-specific survey data, used in refined 1-D MCA
BW	Lognormal	71.75, 14.2	kg	U.S. EPA, 1997a, Tables 7-4 and 7-5; (Combined male/female body weight distributions)

[1]All other exposure parameters are inputted as point estimates (see Table 3-4).
[2]Parameters for lognormal PDF are X ~ Lognormal (arithmetic mean, arithmetic standard deviation) unless otherwise stated. Parameters for triangular PDF are X ~ Triangular (minimum, mode, maximum). Parameters for uniform PDF are X ~ Uniform (minimum, maximum).
[3]A point estimate of 0.189 was used to adjust the surface area skin (SA_skin) distribution, which is based on total body surface area, to account for skin exposures limited to face, forearms, and hands (U.S. EPA, 1997a, Vol. I).
[4]Parameters for preliminary lognormal PDF for ED were converted from a geometric mean of 10 and a 95th percentile of 25.
[5]Parameters for site-specific lognormal PDF for ED are arithmetic mean, standard deviation, and upper truncation limit.

Table 3-7. 1-D MCA Risk Estimates using Preliminary Inputs.

Cumulative Percentile	Spill Area Risk
50th	5.7E-05
90th	1.3E-04
95th	1.6E-04
99th	2.4E-04
99.9th	3.9E-04

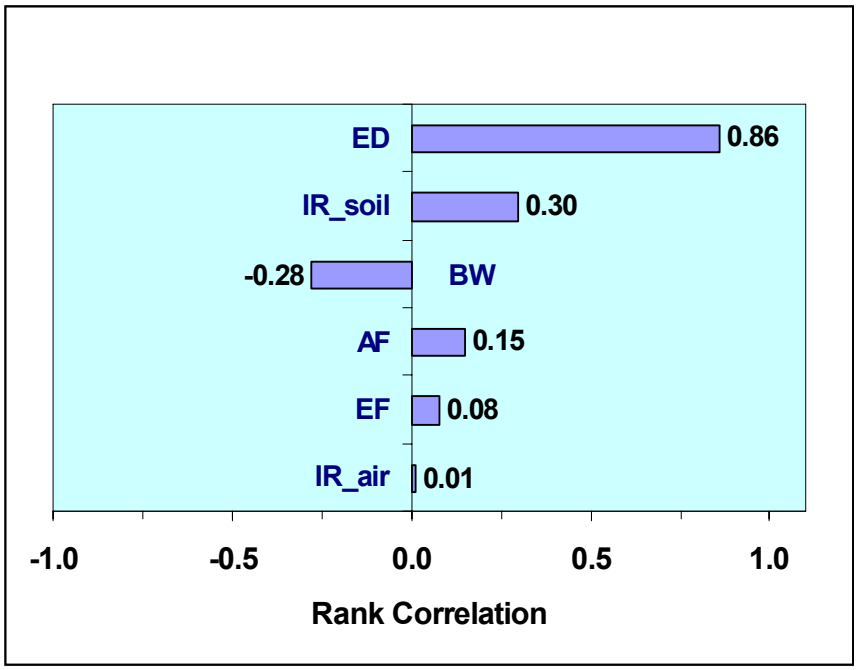


Figure 3-6. Results of sensitivity analysis for preliminary 1-D MCA (Tier 2) showing the Spearman Rank correlations (see Appendix A and B) between input variables and risk estimates.

Table 3-8. Exposure Duration Survey Results.

Survey Results (years)			Summary Statistics	
24.9	20.3	17.2	n	20
8.4	11.7	6.5	min	3.0
3.0	4.7	16.5	max	44.2
6.8	20.9	6.0	arithmetic mean	14.0
18.5	10.6	18.8	standard dev	9.4
9.1	12.7	11.7	median/GM	11.7
7.2	44.2		GSD	1.8

Table 3-9. Refined Point Estimate and 1-D MCA Risk Estimates.

Cumulative Percentile	Spill Area Risk
Refined RME Point Estimate	3.1E-04
50th	6.7E-05
90th	1.6E-04
95th	2.1E-04
99th	3.2E-04
99.9th	5.3E-04

REFERENCES FOR CHAPTER 3

Hedges, L.V. and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. Academic Press, Inc. Orlando.

Oregon DEQ. 1998. *Guidance for the Use of Probabilistic Analysis in Human Health Exposure Assessments*. Waste Management and Cleanup Division. Interim Final. November.

Stiteler, W.M., L.A. Knauf, R.C. Hertzberg, and R.S. Schoeny. 1993. A Statistical Test of Compatibility of Data Sets to a Common Dose-Response Model. *Regulatory Tox. Pharm*. 18: 392–402.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund* (RAGS): *Volume I. Human Health Evaluation Manual* (HHEM) *(Part A, Baseline Risk Assessment).* Interim Final. Office of Emergency and Remedial Response, Washington, DC. EPA/540/1-89/002. NTIS PB90-155581.

U.S. EPA. 1992a. *Final Guidelines for Exposure Assessment*. EPA/600/Z-92/001. 57 *Federal Register*, 22888-22938. May 29.

U.S. EPA. 1992b. *Guidance on Data Usability in Risk Assessment*. Part A. Final. Office of Solid Waste and Emergency Response, Washington, DC. OSWER Directive No. 9285.7.09A. NTIS PB92-96336.

U.S. EPA. 1993. *Data Quality Objectives Process for Superfund*. Office of Solid Waste and Emergency Response. Washington, DC.

U.S. EPA. 1996. *Final Soil Screening Guidance, May 17, 1996. Soil Screening User's Guide*. Office of Solid Waste and Emergency Response, Washington, DC. EPA 540/R-96/018.

U.S. EPA. 1997a. *Exposure Factors Handbook, Volume 1.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fa.

U.S. EPA. 1997b. *Exposure Factors Handbook, Volume 2.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fb.

U.S. EPA. 1997c. *Exposure Factors Handbook, Volume 3.* Office of Research and Development, Washington, DC. EPA/600/P-95/002Fc.

U.S. EPA. 1997d. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment*. Final. National Center for Environmental Assessment, Washington, DC. EPA/630/R-95/002F.

U.S. EPA. 2001. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Office of Solid Waste and Emergency Response. Washington, DC. OSWER Directive No. 9355.4-24. December.

CHAPTER 4

PROBABILISTIC ANALYSIS IN ECOLOGICAL RISK ASSESSMENT

4.1 INTRODUCTION

4.1.1 BASIC APPROACH FOR PERFORMING ECOLOGICAL RISK ASSESSMENTS

Ecological risk assessment (ERA) is defined by the 1997 Environmental Protection Agency's (EPA) *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments (ERAGS)* (U.S. EPA, 1997a) as an evaluation of the "likelihood that adverse ecological effects are occurring or may occur as a result of exposure to one or more stressors". The *ERAGS* document is generally similar to, and consistent with the earlier framework guidance and approach (U.S. EPA, 1992a) which was expanded upon and superceded by the *Guidelines for Ecological Risk Assessment* (U.S. EPA, 1998). The EPA has developed extensive technical and policy guidance on how ERAs should be planned and performed (see Exhibit 4-2). In general, this process has three main elements, as shown in Figure 4-1:

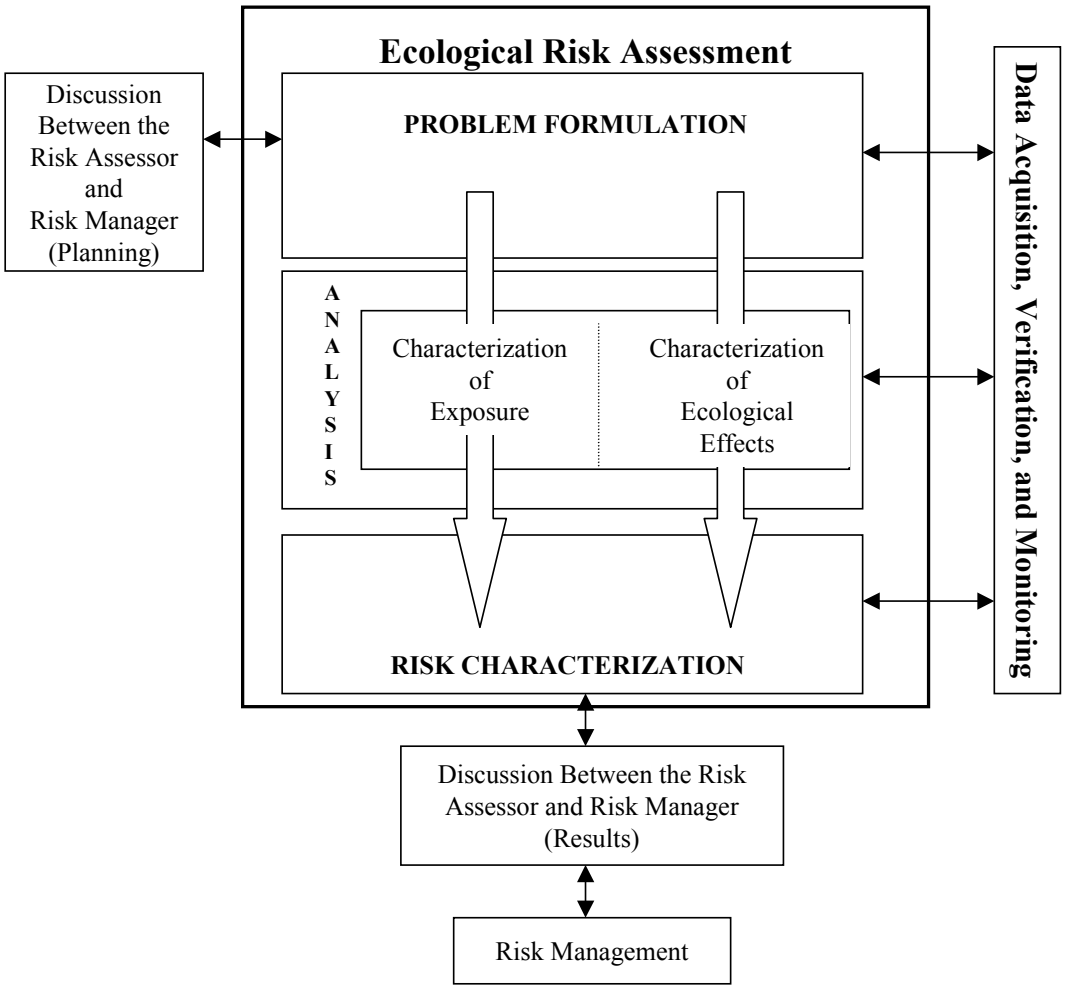


Figure 4-1. Ecological Risk Assessment Framework (U.S. EPA, 1992a)

Problem Formulation provides a foundation for the entire risk assessment. This element includes the specification of risk management goals and assessment endpoints, the development of a site conceptual model with exposure pathways and receptors, and the development of a sampling and analysis plan to collect data on exposures and measures of effects that are needed to support the ERA. In general, problem formulation serves as the foundation of an ERA and often is an iterative process, whereby substantial re-evaluation may occur as new information and data are collected during the site investigations. Collection of data in subsequent iterations is often triggered by identification of major data gaps and uncertainties in the risk characterization that prevent confident decision making by risk managers.

Analysis includes two principal measurement steps that are based upon the problem formulation: Assessment of exposures and assessment of ecological effects. Assessment of exposures includes the identification of stressors at the site that may affect ecological receptors, a characterization of the spatial and/or temporal pattern of the stressors in the environment at the site, and an analysis of the level of contact or co-occurrence between the stressors and the ecological receptors. Assessment of ecological effects includes identification of the types of effects which different stressors may have on ecological receptors, along with a characterization of the relationship between the level of exposure to the stressor and the expected biological or ecological response. This is referred to as the stressor-response relationship.

Risk Characterization combines the exposure characterization and the effects characterization in order to provide a quantitative likelihood or qualitative description of the nature, frequency, and severity of ecological risks attributable to exposure to stressors at a site, as well as an evaluation of the ecological relevance of the effects. Good risk characterizations express results clearly, articulate major assumptions and uncertainties, identify reasonable alternative interpretations, and separate scientific conclusions from policy judgments (U.S. EPA, 1995, 1998).

EXHIBIT 4-1

DEFINITIONS FOR CHAPTER 4

Assessment Endpoint - An explicit expression of an environmental value (ecological resource) that is to be protected, operationally defined by risk managers and risk assessors as valuable attributes of an ecological entity.

Benchmark Dose (BMD) - The dose which causes a specified level of response. The lower confidence limit on the BMD is usually referred to as the BMDL.

Community - An assemblage of populations of different species specified by locales in space and time.

Conceptual Model - A site conceptual model (SCM) in the problem formulation for an ecological risk assessment is a written description and visual representation of predicted relationships between ecological entities and the stressors to which they may be exposed, including sources and pathways of stressors.

Ecological Risk Assessment (ERA) - The process that evaluates the likelihood that adverse ecological effects may occur or are occurring as a result of exposure to one or more stressors.

Lines of Evidence - Information derived from different sources or techniques that can be used to characterize the level of risk posed to exposed receptors; weight-of-evidence generally refers to the quantity of science, while strength of evidence generally refers to the quality of science.

Lowest-Observed-Adverse-Effect Level (LOAEL) - The lowest level of a stressor evaluated in a test that caused a statistically significant effect on one or more measurement endpoints linked to undesirable (adverse) biological changes.

Measurement Endpoint (Measure of Effect) - A measurable ecological property that is related to the valued characteristic chosen as the assessment endpoint. Measurement endpoints (also called measures of effect) often are expressed as the statistical or numeric summaries of the observations that make up the measurement.

No-Observed-Adverse-Effect Level (NOAEL) - The highest level of a stressor administered in a test that did not cause a statistically significant effect in any measurement endpoint linked to an undesirable (adverse) biological change.

Population - An aggregate of individuals of a species within a specified location in space and time.

Receptor - The ecological entity (with various levels of organization) exposed to the stressor.

Risk Characterization (ecological) - The third and last phase of ERA that integrates the analyses of exposure to stressors with associated ecological effects to evaluate likelihoods of adverse ecological effects. The ecological relevance of the adverse effects is discussed, including consideration of the types, severity, and magnitudes of the effects, their spatial and temporal patterns, and the likelihood of recovery.

Scientific/Management Decision Point (SMDP) - A time during the ERA when a risk assessor communicates results or plans of the assessment at that stage to a risk manager. The risk manager decides if information is sufficient to proceed with risk management strategies or whether more information is needed to characterize risk.

Species - A group of organisms that actually or potentially interbreed and are reproductively isolated from similar groups; also, a taxonomic grouping of morphologically similar individuals.

Stressor - Any chemical, physical or biological entity that can induce an adverse response in an ecological receptor; Superfund considers all stressors, but focuses on chemical (toxicant) stressors.

Toxicity Reference Value (TRV) - A dose or concentration used to approximate the exposure threshold for a specified effect in a specified receptor. A TRV is often based on a NOAEL or LOAEL from a laboratory-based test in a relevant receptor species.

EXHIBIT 4-2

ECOLOGICAL RISK ASSESSMENT GUIDANCE AND POLICY DIRECTIVES

EPA has developed extensive guidance and policies on methods and approaches for performing ERAs, including the following:

(1) *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments ("ERAGS"), Interim Final (U.S. EPA, 1997a)*. This document includes processes and steps specifically selected for use in ERAs at Superfund sites. This document supersedes the 1989 *EPA RAGS, Volume II, Environmental Evaluation Manual,* Interim Final (U.S. EPA, 1989). Supplements to ERAGS include the *EcoUpdates* (U.S. EPA, 1991-present, Intermittent Bulletin Series, 1991 to present), which provide brief recommendations on common issues for Superfund ERAs.

(2) *Guidelines for Ecological Risk Assessment ("Guidelines") (U.S. EPA, 1998).* This document updates general (nonprogram specific) guidance that expands upon and replaces the earlier *Framework for Ecological Risk Assessment* (U.S. EPA, 1992a). The approaches and methods outlined in the *Guidelines* and in *ERAGS* are generally consistent with each other.

(3) *Risk Assessment Guidance for Superfund (RAGS): Volume 1–Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments)*, (U.S. EPA, 2001). This guidance specifies formats that are required to present data and results in baseline risk assessments (both human and ecological) at Superfund sites.

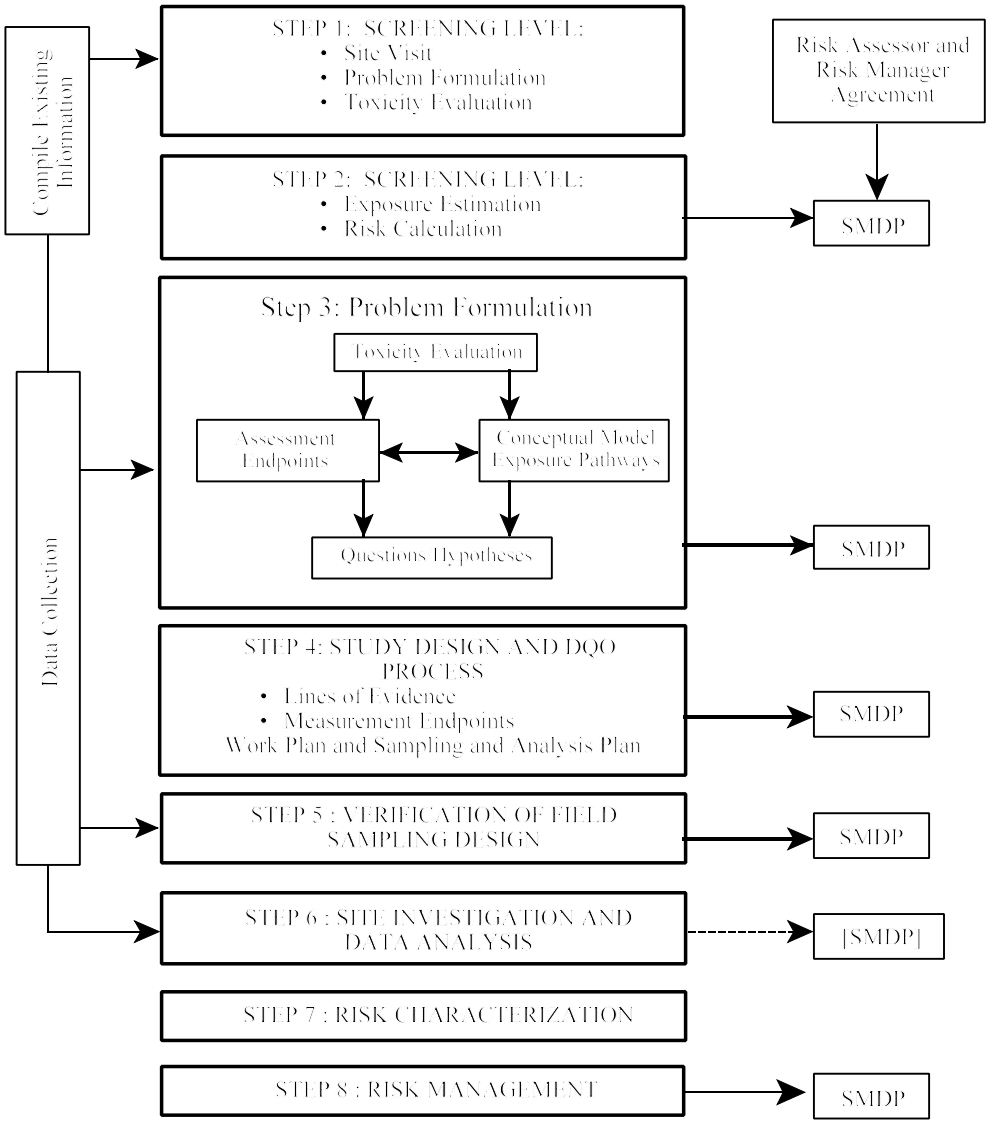
(4) Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors*, F. Henry Habicht, Deputy Administrator, Feb. 26, 1992 (U.S. EPA, 1992b). This policy requires baseline risk assessments to present ranges of risks based on "central tendency" and "reasonable maximum" (RME) or "high-end" exposures with corresponding risk estimates.

(5) Policy Memorandum: *Role of the Ecological Risk Assessment in the Baseline Risk Assessment*, Elliott Laws, Assistant Administrator, August 12, 1994 (U.S. EPA, 1994). This policy requires the same high level of effort and quality for ERAs as commonly performed for human health risk assessments at Superfund sites.

(6) Policy Memorandum: *EPA Risk Characterization Program*, Carol Browner, Administrator, March 21, 1995 (U.S. EPA, 1995). This policy clarifies the presentation of hazards and uncertainty in human and ERAs, calling for clarity, transparency, reasonableness, and consistency.

(7) Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites. Stephen D. Luftig for Larry D. Reed, October 7, 1999 (U.S. EPA, 1999). This document presents six key principles in ecological risk management and decision making at Superfund sites.

ERA is a key component of the remedial investigation process that EPA uses at Superfund sites. *ERAGS* is a program-specific guidance for Superfund that focuses on chemical stressors released into the environment from hazardous waste sites. This guidance refers to ERA as a "qualitative and/or quantitative appraisal of the actual or potential impacts of contaminants from a hazardous waste site on plants and animals other than humans and domesticated species. An excess risk does not exist unless: (1) the stressor has the ability to cause one or more adverse effects, and (2) the stressor co-occurs with or contacts an ecological component long enough and at a sufficient intensity to elicit the identified adverse effect." The *ERAGS* document provides guidance on using an eight-step process for completing an ERA for the Superfund Program, as shown in Figure 4-2.



SMDP= Scientific/Management Decision Point

Figure 4-2. Eight-step Ecological Risk Assessment Process for Superfund (U.S. EPA, 1997a).

4.1.2 PREDICTIVE *VS* OBSERVATIONAL APPROACHES

In general, conclusions about ecological hazards from environmental contamination may be based on information derived from two different techniques: the predictive approach (a comparison of calculated exposures with a set of toxicity reference values), and the observational approach (direct evaluation of the range of potential exposures, coupled with site-specific toxicity testing and population demographic estimates).

Predictive Approach: The core of all Superfund ERAs is the predictive approach, including exposure assessment, toxicity assessment, and risk characterization. The predictive approach is based on a comparison of calculated estimates of chemical exposure of a receptor to one or more Toxicity Reference Values (TRVs) appropriate for that chemical and that receptor. The ratio of exposure at the site to the TRV is referred to as the Hazard Quotient (HQ). The predictive approach has always been used at Superfund sites because it is relatively easy to implement, and because it can be used to evaluate not only current risks, but also risks that might exist in the future if any important changes were to occur in the level of contamination (e.g., due to on-going fate and transport processes), or to changes in land use (a change in land use might alter a number of habitat factors that influence the number and identify of ecological receptors). The predictive approach, however, has the inherent uncertainties of the assumptions in the exposure and toxicity models which are seldom site-specific and thus can lead to either over-protective or under-protective estimates of risk.

Direct Observation: If there is a need to reduce uncertainties in the predictive approach, direct observations of exposure and effects can be collected at Superfund hazardous waste sites. The predictive approach used in ERA does not negate the use of descriptive toxicological approaches or the use of site-specific exposure data, such as toxicity testing or bioaccumulation measurements. Site-specific observations, such as toxicity testing of invertebrates over a gradient of site contaminant exposure levels, may be used to develop site-specific and chemical-specific toxicological relationships. Site-specific measures of exposure or ecosystem characteristics can be used to reduce uncertainty in the exposure assessment and aid in the development of cleanup goals in the Baseline ERA. The direct observation of the exposure and effects on ecological receptors does not however constitute a complete risk assessment. If field or laboratory studies are NOT designed appropriately to elicit stressor-response relationships, direct impacts should not be used as the sole measure of risk because of the difficulty in interpreting and using these results to develop cleanup goals in the ERA. Furthermore, poorly designed toxicological evaluations of environmental media from the site may not allow a definitive identification of the cause of adverse response. For example, receptor abundance and diversity as demographic data reflect many factors (habitat suitability, availability of food, predator-prey relationships among others). If these factors are not properly controlled in the experimental design of the study collecting the observational data, conclusions regarding chemical stressors can be confounded. In addition, direct observation provides information about current risks only and not potential risks should land use or exposure change in the future. Hence, direct observations may be used as a line of evidence in an ERA, but should not be the sole evidence used to characterize the presence or absence of the risks of an adverse effect in the future.

4.1.3 POTENTIAL ADVANTAGES AND LIMITATIONS OF PROBABILISTIC METHODS IN ERA

Probabilistic risk assessment (PRA) is a computational tool that may help increase the strength of the *predictive* evaluation of ecological risks, as well as sometimes helping to better evaluate distributions of observational data for an ERA. The potential advantages of PRA compared to, or possible benefits in augmentation of, the conventional point estimate approach for characterizing variability in exposure or risk are discussed in Chapter 1 and Exhibits 1-6 and 1-7. In brief, point estimate calculations utilize simplifications and assumptions in order to deal with the complex mathematics of combining inputs that are inherently variable. Probabilistic models, in contrast, are designed to combine sets of information on inputs that are expressed as probability distributions. Therefore, PRA generally can yield risk estimates that allow for a more complete characterization of variability and uncertainty, and a potentially more useful sensitivity analysis as compared to estimating sensitivities of inputs from point estimates (see Appendix A). For example, sensitivity analysis can help determine major contributors to exposure factors and sources of uncertainty that could help to design better sampling and analysis plans in later iterations to help fill data gaps and reduce uncertainties for risk characterization.

Because of the inherent differences in the computational approach, as in the case with any additional risk assessment information, PRA may sometimes lead to a different risk assessment outcome and risk management decision than would be derived from the use of point estimate calculations alone. The differences in the decisions stemming from the two approaches will vary from case to case, depending mainly on the form of the exposure or risk model, the attributes of the distributions of the input values, and the quality, quantity, and representativeness of the data on which the input distributions are derived. Sometimes the differences between the two approaches will be quite large, and the information gained from a PRA can play an important role as weight-of-evidence in communicating risks to stakeholders and risk managers.

Even though PRA may have some advantages, it also has limitations and potential for misuse. PRA can not fill basic data gaps and can not eliminate all of the potential concerns associated with those data gaps. That is, if one or more of the input distributions are not well characterized and the distribution(s) must be estimated or assumed, then the results of the PRA approach will share the same uncertainty as the point estimate values. However, given equal states of knowledge, the PRA approach may yield a more complete characterization of the exposure or risk distribution than the point estimate approach.

Of course, any prediction of exposure or risk is based on the use of mathematical models to represent very complex environmental, biological, and ecological systems. No matter how sophisticated the computations, questions will always exist as to whether the calculated values are a good approximation of the truth. Therefore, even when PRA is used as a supplemental tool to point estimations (deterministic) of risks in the ERA process, a weight-of-evidence approach that combines the predictive approach with direct observations will still provide the most appropriate basis for decision making.

A second application of PRA in ERA, besides the characterization and incorporation of distributions of data for ERA, is the characterization of uncertainty in calculated estimates of exposure or risk. In this application, whatever uncertainty may exist in one or more of the input distributions is characterized, and quantitative estimates of the confidence limits around the mean, upper bound, or any other percentile of the output distribution are calculated. This use of PRA is often especially important in risk management decision making, since the range of uncertainty around central tendency exposure (CTE) and reasonable maximum exposure (RME) or other upper bound estimates of exposure or risk can

sometimes be quite large. As stated before, the point estimate approach can also provide estimates of uncertainty, but the PRA approach often provides a more complete characterization of the uncertainty.

4.1.4 FOCUS OF THIS CHAPTER

This chapter focuses on the application of PRA as a tool for predicting ecological risks at Superfund sites. Some of the methods and approaches described in this chapter are similar to those that have been developed by U.S. EPA's Office of Pesticide Programs Committee on Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) Risk Assessment Methods (ECOFRAM, 1999a, 1999b) for use in assessing environmental hazards of pesticide products. However, the methods described in this chapter are specifically designed to be applicable at Superfund sites and to be consistent with other Superfund guidance.

This chapter does not seek to provide guidance on the many basic issues that must be faced in planning and performing any ERA. Prior to considering the use of PRA in an ERA, fundamental concepts will already have been developed, such as a problem formulation with a conceptual site model, selection of representative receptors, definition of exposed populations, definition of risk management objectives and goals, selection of assessment endpoints, calculation of TRVs and development of site sampling plans, etc. Likewise, this chapter does not repeat the presentation of basic statistical and mathematical methods used in PRA, since these are described in other chapters and appendices of this document. In summary:

- ☞ *This chapter focuses on application of PRA techniques to ERA at Superfund sites.*
- ☞ *The reader is assumed to be familiar with the basic methods used in ERA at Superfund sites, and this chapter does not address basic tactical and technical issues in ERA.*
- ☞ *The reader is assumed to be familiar with the basic mathematical principles and techniques of PRA as described in other chapters and appendices of this document.*

4.2 DECIDING IF AND WHEN TO USE PRA IN ECOLOGICAL RISK ASSESSMENT

As shown in Figure 4-2, the ERA process for Superfund includes a number of scientific/ management decision points (SMDPs) (U.S. EPA, 1997a). The SMDP is a point of consultation between the risk manager, EPA Regional Biological Technical Assistance Group (BTAG) coordinator, EPA regional ecotoxicologist, and other stakeholders, and is intended to provide an opportunity for re-evaluation of direction and goals of the assessment at critical points in the process. It is during the SMDP discussions that it is important to decide whether or not a PRA is likely to be useful in decision making. If so, the pursuit of distributed data is justified. Within the 8-step process of developing the ERA, PRA could provide insight at several steps. A decision to move forward with distributional analyses should be considered within the BTAG context during the documentation of the outcome of the SMDPs after Step 3 within the process. As a reminder, PRA is NOT intended to be a replacement for point estimate analyses; rather PRA supplements the required presentation of point estimates of risk. It is also emphasized that the use of PRA should never be viewed as or used in an attempt to simply generate an alternative risk estimate or PRG, compared to that which was derived by a point estimate ERA; instead, PRA should be

used to provide insightful information on distributions of various factors (exposure, toxicity, and hazards) which can provide weight-of-evidence evaluations of potential risks in conjunction with a point estimate ERA. There are a number of factors to consider in making these decisions, as discussed below.

4.2.1 TECHNICAL CONSIDERATIONS

The fundamental reason for performing any predictive risk assessment (point estimate or probabilistic) is to provide information to risk managers in order to help support the risk management decision-making process. As noted above, a properly performed PRA may help to yield more description of variability in exposure and risk than can be achieved using the point estimate approach. Therefore, if any of a site's data may be better described and evaluated by distributions, then a PRA can be applied to any part of an ERA or even to the entire ERA for expressing risk characterization in probabilistic terms; again, always in conjunction with the required point estimate ERA. However, when risk estimates derived from the point estimate approach are either far below or far above a level of risk management concern, any such potential improvements in risk characterization are not likely to influence risk management decision making. In these cases, PRA is not likely to be as useful in decision making. Even so, PRA may help in these situations by providing information that may be useful in better deciding where the gradient of excess risks are reduced to acceptable levels. Rather, it is more common for a PRA to be useful when point estimates of risks are close to the decision threshold (such that PRA-based refinements in the risk estimates might be important in making risk management decisions). It is for this reason that PRA may be useful to apply either during the development of the ERA after the screen (Steps 3 to 6, U.S. EPA, 1997a), or after point estimate results from the baseline ERA have been completed (Steps 1 to 7, U.S. EPA, 1997a).

The results of a point estimate risk assessment will normally present the range of risks based on central tendency exposure and reasonable maximum exposure input assumptions and on the no-observed-adverse-effect-level (NOAEL)- and lowest-observed-adverse-effect-level (LOAEL)-based TRVs (U.S. EPA, 1992b, 1997b). The bounds for the highest HQ are derived from the ratio of the RME compared to the NOAEL-based TRV, and the bounds for the lowest HQ are based on the ratio of the CTE compared to the LOAEL-based TRV. These two bounded extreme estimates of risk can be used to screen out cases where PRA is not likely to be as useful. That is, if the risk to the RME receptor is clearly below a level of concern using the NOAEL-based TRV, then risks to the exposed population are likely to be low and PRA analysis is likely not needed. Likewise, if risks to the CTE receptor are clearly above a level of concern using the LOAEL-based TRV, then risks to the exposed population are likely to be of definite concern, and a PRA may not provide as much additional useful information to the risk manager, except in the case where uncertainties remain high and the derivation of an appropriate and realistic clean-up goal may be difficult. If the risks are intermediate between these two bounds (e.g., risks to the CTE receptor are below a level of concern based on the LOAEL-based TRV but are above a level of concern based on the NOAEL-based TRV), then PRA might be helpful in further characterizing the site risks in balance with the point estimates of risks and in supporting decision making or in deciding if additional iterations of analyses would be needed. This concept is illustrated graphically in Figure 4-3.

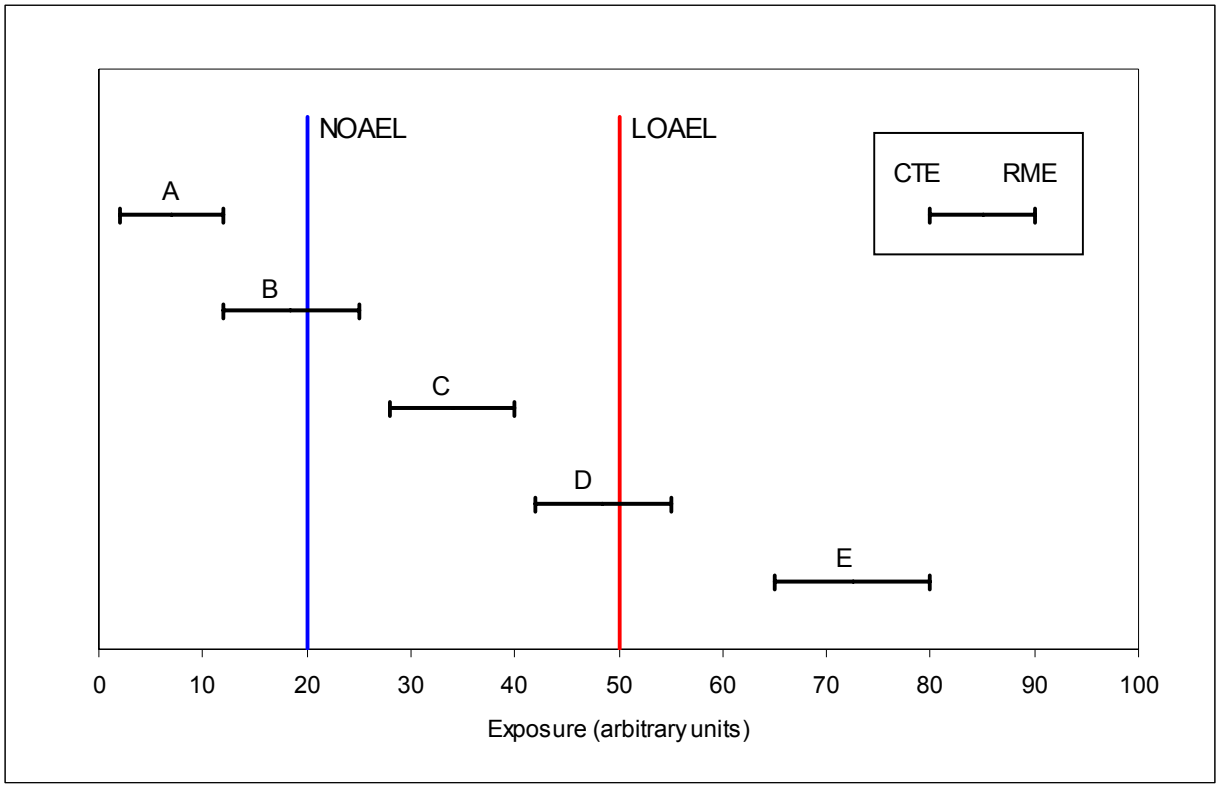


Figure 4-3. Example of cases where use of PRA may be helpful. In cases A and E, the range of risks (CTE to RME) estimated by the point estimate method are either well below (Case A) or well above (Case E) the likely level of concern based on the NOAEL-LOAEL range, and PRA is not likely to alter risk management decisions regarding the potential need for remediation. In cases B, C, and D, the point estimates of risk overlap or fall within the range of potential concern, suggesting that PRA-based risk estimates might be helpful in supporting risk management decisions.

The second main technical reason to consider conducting PRA is that the PRA methodology can help characterize and quantify the degree of variability and uncertainty around any particular estimate of exposure or risk (e.g., the CTE or RME). The purpose of the analysis would be to estimate the uncertainty around an exposure or toxicity or risk estimate, generally with little or no additional data acquisition. The only additional information needed to perform the analysis is an estimate of the uncertainty in the true parameter values of the key variables in the variability model. In some cases, these estimates of uncertainty around parameter values may be developed from statistical analysis of the available data. Alternatively, professional judgment may be used to establish credible bounds on the parameters, especially when relevant data are sparse.

☞ *Even in the presence of data gaps, uncertainty analysis using PRA can provide useful information. Indeed, it is when data are limiting or absent that a quantitative probabilistic analysis of uncertainty may be most helpful.*

4.2.2 COST AND SCHEDULE CONSIDERATIONS

Performing a PRA can sometimes add time and cost to an ERA. As discussed in Chapter 2, in part, the decision to progress from a point estimate assessment to a PRA reflects a belief that the potential value of the PRA for risk management decision making outweighs the additional time and costs. The tiered process encourages a systematic approach for both the point estimate and probabilistic assessments, whereby the least complex methods are applied first. For example, the initial Tier 2 assessment may be conducted with a set of preliminary probability distributions for variability (PDFv), developed with much the same information and assumptions that were applied to develop point estimates in Tier 1. Parameter values can be estimated by setting the arithmetic mean equal to the CTE point estimate, and the 95th percentile equal to the RME point estimate. The choice of distributions may differ depending on the state of knowledge for a particular variable (see Appendix B). For example, unbounded variables might be characterized with lognormal distributions while bounded distributions are characterized by beta or Johnson Sb distributions. Certain variables may continue to be characterized by point estimates, especially if the sensitivity analysis suggests that the chemical, pathway, and/or exposure variables are relatively minor contributors to total exposure and risk. The decision to collect additional data or explore alternative methods for developing probability distributions can be reexamined in an iterative fashion by evaluating the expected benefits of the added information to the risk management decision-making process. These concepts are presented in greater detail in Chapter 2 (see Figures 2-1 and 2-2).

4.3 PROBLEM FORMULATION

Once a decision has been made to include PRA in an ERA, the first step should be to re-visit the problem formulation step and carefully determine the scope and objectives of the PRA. Typically, a considerable amount of knowledge will have been gained during the screening level and baseline point estimate evaluations, and this knowledge should be used to help focus and narrow the scope of the PRA. That is, the PRA will generally utilize the same basic exposure and risk models used in the point estimate approach, but the PRA will typically evaluate only a sub-set of the scenarios considered. For example, chemicals, pathways, and/or receptors that are found to contribute a negligible level of exposure or risk may usually be omitted from the PRA, while those factors that contribute significantly to an excess level of risk concern in the point estimate approach should generally be retained. As noted previously, when a chemical or pathway is omitted from a PRA analysis, this does not mean that it is eliminated from the overall risk assessment; rather, it may be kept in the assessment as a point estimate.

The next step in problem formulation for a PRA should be to define whether the goal of the analysis is to characterize variability alone, or to characterize both variability and uncertainty. In either case, sensitivity analysis (as summarized in the preceding paragraph, or for more details see Appendix A) should be used to help identify which of the input variables contribute the most to the variability in the outputs (exposure, toxic effects, or risk), and the initial PRA should focus on defining the probability density functions (PDFs) for those input variables. An analysis of uncertainty, if thought to provide additional useful information, may also be included at the initial level, or may be delayed until the initial analysis of variability is completed.

As always, problem formulation should be viewed as an iterative process, and it is reasonable and appropriate that decisions regarding the scope and direction of the PRA should be reassessed (at SMDPs) as information becomes available from the initial evaluations. As stressed above, the fundamental criterion which should be used is whether or not further PRA evaluations are likely to provide additional information to a point estimate ERA that will help strengthen and support the risk management decision-making process.

4.4 MODELING VARIABILITY IN EXPOSURE

There are two main types of descriptors of exposure that may be used in ERA: dose and concentration. For terrestrial receptors such as mammals or birds, exposure is most often described in terms of ingested dose (mg/kg-day). In most cases, this will be based on chemical ingested from drinking water and/or the diet, including incidental soil ingestion, but could also include amounts of chemical taken up across the skin or through inhalation as additional routes of exposure. The exposure levels are most often expressed as doses, since that term tends to normalize the confounding factors of variable daily intake rates and body weights that occur if/when one only evaluates concentrations. For aquatic receptors, the main route of exposure is usually by direct contact and less often by ingestion, so exposure is usually characterized in terms of concentration of contaminants in surface water, pore water and/or sediment. Likewise, exposure of terrestrial plants and terrestrial invertebrates, such as earthworms, is usually described in terms of concentration of contaminants in soil. In some cases, exposure of terrestrial receptors is characterized in terms of specific tissue or whole-body concentrations of contaminants. Examples of calculating and presenting dose-based and concentration-based distributions of exposure are presented below.

4.4.1 CHARACTERIZING VARIABILITY IN DOSE

The general equation used for calculating the **dose** of a contaminant of concern in a specified environmental medium (e.g., water, soil, air, diet, etc.) by a particular member of a population of exposed receptors is:

$DI_{i,j} = C_i \times IR_{i,j} / BW_j$

where:

$DI_{i,j}$ = Average daily intake of chemical due to ingestion of medium "I" by a population member "j" of the exposed population (mg/kg-day)

C_i = Concentration of chemical in environmental medium "I" (mg/unit medium)

$IR_{i,j}$ = Intake rate of medium "I" at the site by population member "j" (units of medium per day)

BW_j = Body weight of population member "j" (kg)

Total exposure of a population member "j" is then the sum of the exposures across the different media:

$DI_{total,j} = \sum DI_{i,j}$

In this basic equation, $IR_{i,j}$ and BW_j are random variables (i.e., they have different measurable values for different members of the exposed population) that are often correlated. For example, a receptor with a relatively low intake rate can also be expected to have a low body weight. Some studies utilize paired measurements of IR and BW by individual, and present a distribution of the ratio ($IR_{i,j}$ /BW_j), referred to as a body weight-normalized intake rate (mg/kg-day). This expression provides an alternative to using a correlation coefficient to relate two input variables (see Appendix B), and can be entered into the dose equation as follows:

$$DI_{i,j} = C_i \times \left(\frac{IR_{i,j}}{BW_j} \right)$$

where the ratio is characterized by a single probability distribution. Because the variability in this ratio is likely to be different than the variability in the ratio of the IR and BW variables treated independently,

accounting for the correlation can affect the distribution of dose and risk. If empirical data for quantifying the ratio are limited but a relationship is expected, plausible ranges of correlations may be explored as a source of uncertainty in the risk estimates.

The concentration term (C_i) may be characterized by a point estimate or a probability distribution, depending on the relationship between the geographic scales of the measurement data and receptor home range (see Appendix C, Section C.3.1). If the home range of the receptor is small compared to the spatial distribution of sampling locations, C_i may be characterized by the probability distribution for variability in measured concentrations. Alternatively, if the home range is large compared with the exposure area evaluated, then a point estimate (e.g., mean or uncertainty in the mean) may be more appropriate.

In the PRA approach, PDFs should be defined for as many of the input variables as reasonable, especially for those variables that are judged (via sensitivity analysis) to contribute the most to the variability in total exposure. The basic principles for selecting the key variables to model as PDFs are presented in Appendix A, and the basic methods used for selecting and fitting distributions are described in detail in Appendix B.

Figure 4-4 shows several examples of graphical formats which may be used to present the estimated distribution of ingested doses in an exposed population. If a single distribution is plotted (top panel), the PDF format is usually the most familiar and useful for risk assessors and managers, but the cumulative distribution function (CDF) format tends to be less cluttered when multiple distributions are shown (e.g., compare the middle graph to the bottom graph). In addition, percentiles can be read directly from a CDF format, but not from a PDF format graph. In all cases, it is very useful to superimpose the CTE and RME point estimate ranges of exposure directly on the same graph as is used to show the distribution of exposures estimated by PRA. This provides a convenient way to compare the results of the two alternative computational methods, and interpret additional information that the PRA can add to the point estimate ERA.

☞ *A conventional point estimate, range of exposure (CTE to RME) or toxicity (NOAEL to LOAEL) and corresponding risk ranges should be calculated and presented for comparison with the PRA results.*

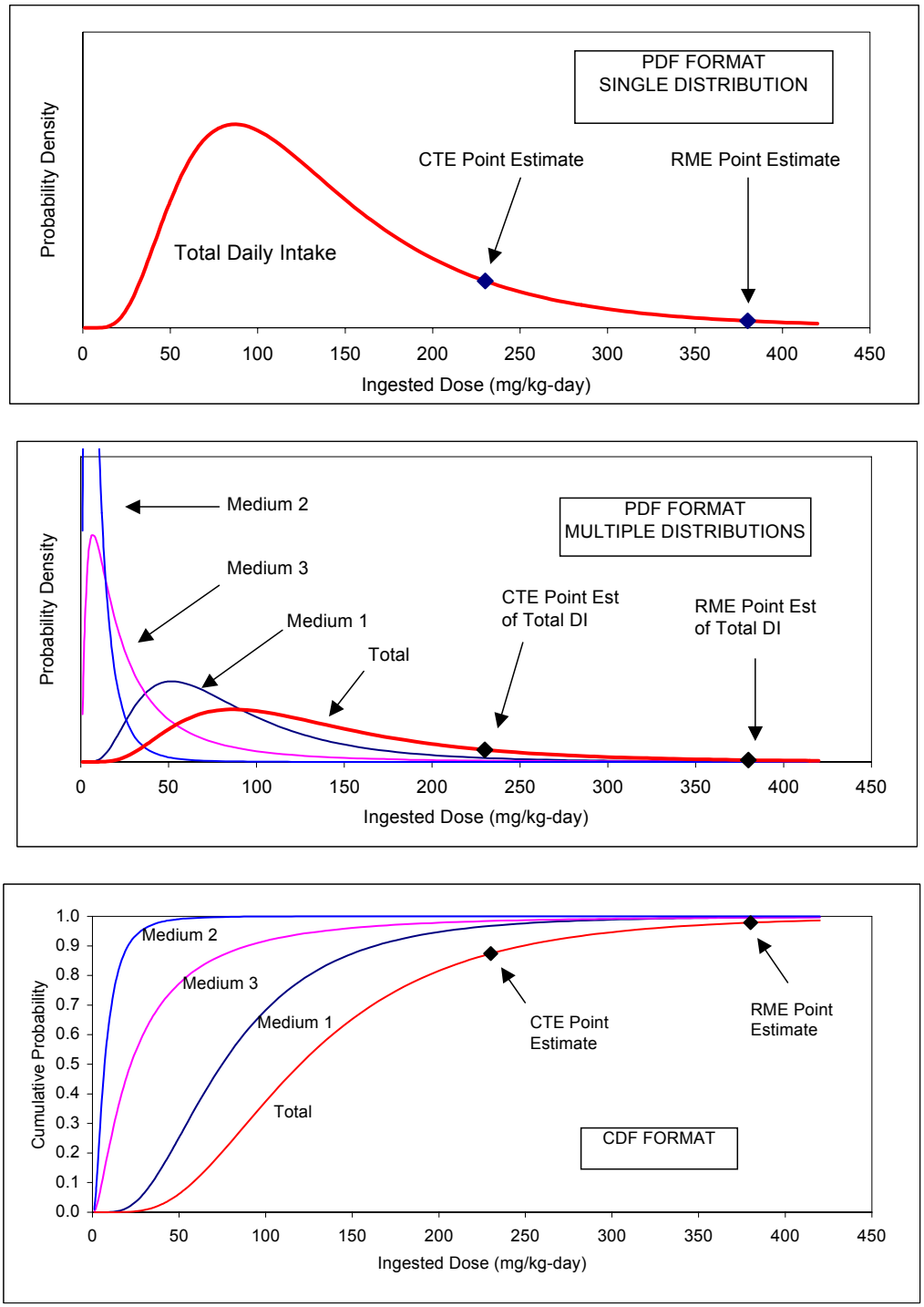


Figure 4-4. Example Graphical Presentations of Dose Distributions.

4.4.2 CHARACTERIZING VARIABILITY IN EXPOSURE CONCENTRATION

As noted above, in some cases the most appropriate descriptor of exposure is concentration (either in an abiotic medium such as water, soil, or sediment, or in the tissues of the receptor), rather than ingested dose. Assuming that the concentration values in the medium of concern are measured rather than modeled, PRA is not required to generate the distribution of concentrations. Rather, the available data may be used to define an appropriate theoretical or empirical distribution function (EDF), as described in Appendix B. If concentrations in the medium are modeled (calculated by PRA) rather than measured, then the exposure distribution may be estimated by using distribution functions (PDFs or CDFs, rather than using point estimates as inputs to the fate and transport model(s) and/or uptake models that predict the concentration levels in the medium of concern. The resulting distribution(s) of concentration may be displayed graphically using the same formats as illustrated in Figure 4-4, except that the x-axis has units of concentration rather than dose. As above, the point estimate ranges of concentration used in the CTE and RME calculations should be plotted on the same graphs to provide a convenient basis for comparing the results of the two approaches and to help interpret the additional information that the PRA can add to the point estimate outputs.

4.5 MODELING VARIABILITY IN TOXICITY

4.5.1 VARIABILITY IN RESPONSE AMONG MEMBERS OF A POPULATION

Data on the toxicity of a chemical usually comes from laboratory studies whereby groups of organisms (laboratory mammals, fish, benthic organisms, plants, earthworms, etc.) are exposed to differing levels of chemical, and one or more responses (endpoints) are measured (survival, growth, reproduction, etc.). These toxicological observations define the exposure-based stressor-response curve that is characteristic for that specific receptor, chemical, and response.

In the point estimate approach, information from the dose/stressor-response curve is generally converted to one or more TRVs, each representing a specific point on the dose-based or concentration-based stressor-response curve. For example, the highest dose or concentration that did not cause a statistically significant change in a toxicologically significant endpoint is defined as either the NOAEL dose or the no-observed-effect concentration (NOEC), while the lowest dose or concentration that did cause a statistically significant effect on a relevant endpoint is the LOAEL dose or the lowest-observed-effect concentration (LOEC). Generally, exposures below NOAEL- or NOEC-based TRVs are interpreted to pose acceptable risk, while exposures above LOAEL- or LOEC-based exposures are judged to pose potentially unacceptable risk. It is essential to note the need for high quality toxicity data to derive reliable and confident TRVs. Strong sampling and study designs, that generate data for site exposure factors and toxicological stressor-response relationships, are of critical importance for producing high quality ERAs by either point estimate or PRA approaches. Shortcomings in either area could be major data gaps or uncertainties that detract from the confidence in the risk characterization of the ERA, and may be a basis for pursuing additional iterations of sampling or studies that are more strongly designed to fill those critical data gaps and reduce uncertainty.

Use of the TRV approach, however, does have some potential limitations. Most important is that the ability of a study to detect an adverse effect depends on both the range of doses tested and the statistical power of the study (i.e., the ability to detect an effect if it occurs). Thus, studies with low power (e.g., those with only a few test animals per dose group) tend to yield NOAEL or NOEC values that are higher than studies with good power (those with many animals per dose group). In addition, the choice of the TRV is restricted to doses or concentrations that were tested, which may or may not be close

to the true threshold for adverse effects, and this uncertainty increases as the interval between doses increases. Finally, it is not always easy to interpret the significance of an exposure that exceeds some particular TRV, since the severity and incidence of response depends on the shape and slope of the exposure response curve (information that is not captured in a point estimate TRV).

One way to resolve some of these stressor-response limitations is to apply uncertainty factors to the NOAEL or NOEC and LOAEL or LOEC, which calculates an adjusted TRV that reduces the study's exposure level of concern to account for those uncertainties, so that there is a lesser chance of overlooking possible adverse exposures (i.e., avoiding a false negative conclusion). Another way to resolve some of the stressor-response limitations is to fit a mathematical equation to the available exposure-response data and describe the entire exposure-response curve. This may be done using any convenient data fitting software, but EPA has developed a software package specifically designed for this type of effort. This software is referred to as the Benchmark Dose Software (BMDS), and is available along with detailed documentation and explanation of the methodology at **www.epa.gov/ncea/bmds.htm**.

The most appropriate mathematical form of the exposure-response model depends on whether the endpoint measured is discrete and dichotomous (e.g., survival) or continuous (e.g., growth rate). For a dichotomous endpoint, the result of the fitting exercise is a mathematical exposure-response model P that yields the probability of a response in an individual exposed at any specified level of exposure (expressed either as dose or concentration). Exhibit 4-3 shows an example of this process using hypothetical data. Thus, for an individual with an exposure level of "x", the probability of a response in that individual is simply P(x). In a population of individuals with exposures x1, x2, x3, ...xi, the expected number of responses (e.g., deaths) in the exposed population is the sum of the probabilities across all individuals in the population. Stated another way, the average fraction of the population that will experience the response is given by the expected value of P (i.e., the average value of P(x)).

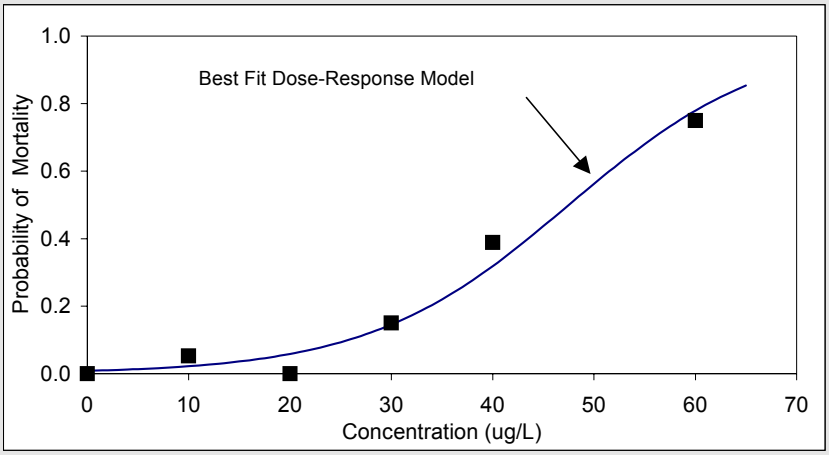
EXHIBIT 4-3

MODELING VARIABILITY IN RESPONSE FOR A DICHOTOMOUS ENDPOINT

The following data are from a hypothetical study of the acute lethality (24 hour) of a chemical using fathead minnows as the test organism:

Concentration ug/L	Number Tested	Survival	
		Dead	Alive
0	20	0	20
10	19	1	18
20	20	0	20
30	20	3	17
40	18	7	11
60	20	15	5

These data were fit to each of the dichotomous models available in BMDS. The best-fit model was the logistic equation. A graph of the best fit curve is shown below.



Basic Equation

Probability of mortality (conc) = 1 / (1 + exp(-a - b*conc))

Best fit parameters

a	-4.80
b	0.101

Goodness of Fit

P	0.604	P=Chi Square Goodness of Fit test statistic
AIC	79.12	AIC=Akaike's Information Criterion

For a continuous endpoint, the BMDS software yields equations that give the expected mean response m(x) at a specified exposure level, along with the standard deviation s(x) that characterizes how variable the response is among different individuals exposed at that same exposure level. The standard deviation may be modeled either as a constant (homogeneous variance) or a function of the exposure level (heterogeneous variance), with the choice depending on which approach yields the best agreement with the observed variances. In most cases there will not be sufficient data to allow a meaningful analysis of the true shape of the underlying distribution of responses at a given exposure, so the choice of the distributional form of the variability in response will require an assumption. In the absence of any clear evidence to the contrary, it is considered likely that the distribution of responses will not be strongly skewed, and that the distribution may be reasonably well modeled using a normal PDF (truncated as necessary to prohibit selection of biologically impossible or implausible values). Thus, variability in response at dose "x" may generally be modeled as:

Response(x) ~ NORMAL[m(x), s(x), min, max]

However, if available data suggest some other distributional form is more appropriate, that form should be used and justified.

Exhibit 4-4 shows an example of this process using hypothetical data. In this case, the mean response was found to be well modeled by the Hill equation, and the standard deviation was found to be best characterized as a constant (rho=0). Thus, given an exposure level "x", the mean response m(x) may be calculated from the model, and this value along with the standard deviation may then be used as parameters for an appropriate type of PDF (e.g., normal) to describe the expected distribution of responses in a population of different individuals exposed at level "x". Section 4.7.2 describes methods that may be used to characterize and quantify the uncertainty associated with this approach.

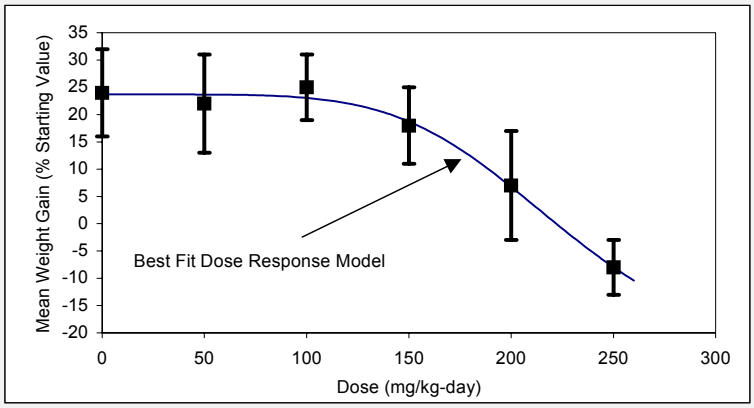
EXHIBIT 4-4

MODELING VARIABILITY IN RESPONSE FOR A CONTINUOUS ENDPOINT

The following data are from a hypothetical study of the effects of a chemical on the growth of laboratory mice. Animals were exposed to the chemical via drinking water for 21 days. The measurement endpoint was weight gain, expressed as a percentage of the starting weight of each animal.

Ingested dose mg/kg-day	Number Tested	Weight Gain (% Starting Value)	
		Mean	Stdev
0	5	24	8
50	5	22	9
100	5	25	6
150	5	18	7
200	5	7	10
250	5	-8	5

These data were fit to each of the continuous models available in BMDS. The best-fit model was the Hill equation with constant variance. A graph of the best fit curve is shown below.



Basic Equations

Mean Response(d) = int + v*d^n / (k^n + d^n)
Variance(d) = alpha*mean response(d)^rho

Best fit parameters

int	23.70
v	-51.41
n	5.295
k	228.7
alpha	48.5
rho	0 (constant variance)

Goodness of Fit

P	0.685	P=Chi Square Goodness of Fit test statistic
AIC	154.5	AIC=Akaike's Information Criterion

4.5.2 VARIABILITY IN RESPONSE AMONG SPECIES

In some cases, risk management decisions may also consider community-level effects as well as population-level or sub-populations effects. That is, a stressor might be considered to be below a level of concern for the sustainability of a community if only a small fraction of the total number of exposed species are affected. In this case, toxicological responses may be best characterized by the distribution of toxicity values across species. This is referred to as a Species Sensitivity Distribution (SSD). This type of approach is generally used for communities of aquatic receptors, since all of the different species that make up the community (e.g., all fish, benthic invertebrates, aquatic plants, and amphibians that reside in a stream) will be exposed to approximately the same concentration of contaminant in the water. The process for generating an SSD consists of the following steps:

(1) Select an appropriate type of endpoint (lethality, growth, reproduction, etc.), and select an appropriate type of point estimate from the exposure-response curve for each species. For example, the TRV might be the LC50 for lethality or the EC20 for growth. The key requirement is that the SSD be composed of TRV endpoints that are all of the same type, not a mixture.

(2) Collect all reliable values for that type of TRV from the literature for as many relevant species as possible. When more than one value is available for a particular species, either select the value that is judged to be of highest quality and/or highest relevance, or combine the values across studies to derive a single composite TRV for each species. It is important to have only one value per species to maintain equal weighting across species.

(3) Characterize the distribution of TRVs across species with an appropriate CDF. Note that there is no *a priori* reason to expect that an SSD will be well characterized by a parametric distribution, so both parametric and empirical distributions should be considered.

Once an SSD has been developed, the fraction of species in the exposed community that may be affected at some specified concentration may be determined either from the empirical distribution or from the fitted distribution. Exhibit 4-5 shows examples of this approach. In this hypothetical case, the TRV selected for use was the LC_{low} (in this case, the LC_{low} is defined as all LC values <=LC10). A total of 13 such values were located. The first graphical presentation is the empirical distribution function, where the Rank Order Statistic (ROS) of each value is plotted as a function of the log of the corresponding value. This may be used directly to estimate the fraction of the species in a community that will be affected by any particular environmental concentration. For example, in this case, it may be seen that a concentration of 10 ug/L would be expected to exceed the LC_{low} for about 33% of the aquatic species for which toxicity data are available. The second graph shows how the data may be characterized by fitting to a continuous distribution. In this case, a lognormal distribution was selected as a matter of convenience, but other distributions may also yield acceptable fits. Based on the best fit lognormal distribution for the SSD data, it is calculated that a concentration of 10 ug/L would be expected to impact about 31% of the exposed species. However, as noted above, there is no special reason to expect that an SSD will be well characterized by a continuous parametric distribution, so some caution should be used in the use of a continuous distribution to fit an SSD, especially when the SSD is based on a limited number of species and when the purpose of the SSD is to estimate percentiles and exposures outside the observed range. The risk assessor should always present an evaluation of the robustness of an SSD to aid in the decision process.

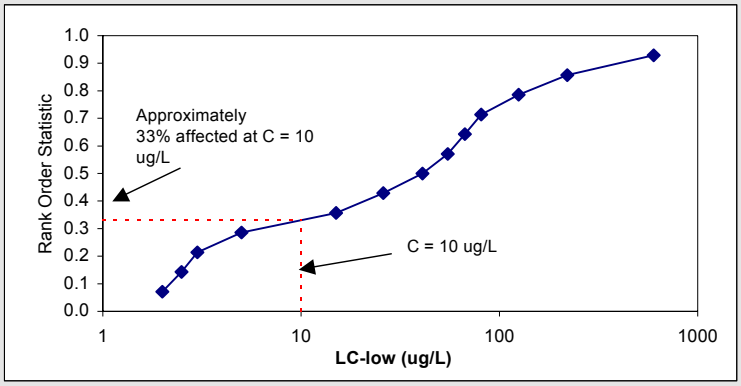
EXHIBIT 4-5

HYPOTHETICAL SPECIES SENSITIVITY DISTRIBUTION

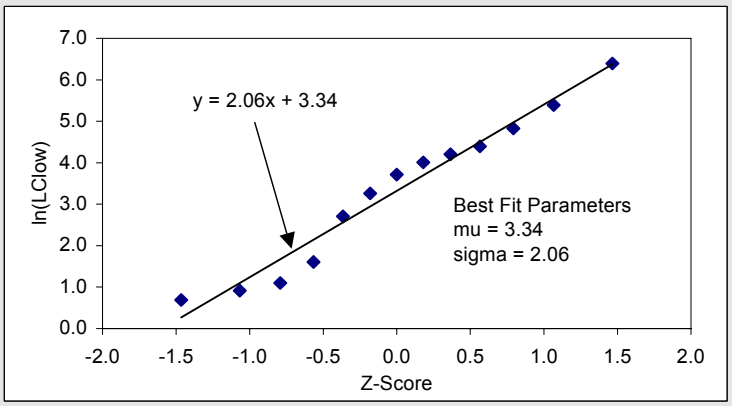
Hypothetical Data

Species	LC_{low}	$ln(LC_{low})$	Rank	ROS	z-score
a	2	0.693	1	0.07	-1.465
b	2.5	0.916	2	0.14	-1.068
c	3	1.099	3	0.21	-0.792
d	5	1.609	4	0.29	-0.566
e	15	2.708	5	0.36	-0.366
f	26	3.258	6	0.43	-0.180
g	41	3.714	7	0.50	0.000
h	55	4.007	8	0.57	0.180
i	67	4.205	9	0.64	0.366
j	81	4.394	10	0.71	0.566
k	125	4.828	11	0.79	0.792
l	220	5.394	12	0.86	1.068
m	600	6.397	13	0.93	1.465

Example EDF: ROS vs LC_{low} (log-scale)



Example Parametric Fit: (Lognormal)



4.6 MODELING VARIABILITY IN RISK

4.6.1 VARIABILITY IN HAZARD QUOTIENT

As noted above, the most common descriptor of risk used in predictive risk assessments is the Hazard Quotient (HQ). The HQ is the ratio of the exposure for some generalized or typical hypothetical member of the receptor population at a site, compared to an appropriate TRV value that equates to an acceptable level of risk for that receptor and chemical. Usually the HQ approach is not based on a single value, but on a range of values in which different levels of exposure (CTE and RME) are compared to both the NOAEL to LOAEL benchmarks. In general, HQ values below 1 are interpreted as indicating acceptable risk, while HQ values above 1 are interpreted as indicating the potential for adverse effects.

Because exposure varies among different members of an exposed population of receptors, HQ values also vary among members of the exposed population. Several alternative approaches for characterizing this variability by PRA methods are presented below.

Variability Within a Population

Figure 4-5 illustrates the simplest approach for summarizing variability in HQ values among the members of an exposed population. In this format, the TRV values appropriate for a particular exposure are simply superimposed on the graph illustrating the distribution of exposures. This may be done either for a dose-based (as shown in the figure) or for a concentration-based exposure parameter. This format allows an easy evaluation of the fraction of the population above (HQ > 1) and below (HQ < 1) each TRV, especially when presented in CDF format. However, this format does not allow for a quantitative estimate of the fraction of the population with HQ values above any value other than 1, although a similar calculation and presentation could be made for any multiple of the TRVs, which would directly equate to that multiple of the HQ (e.g., depicting the

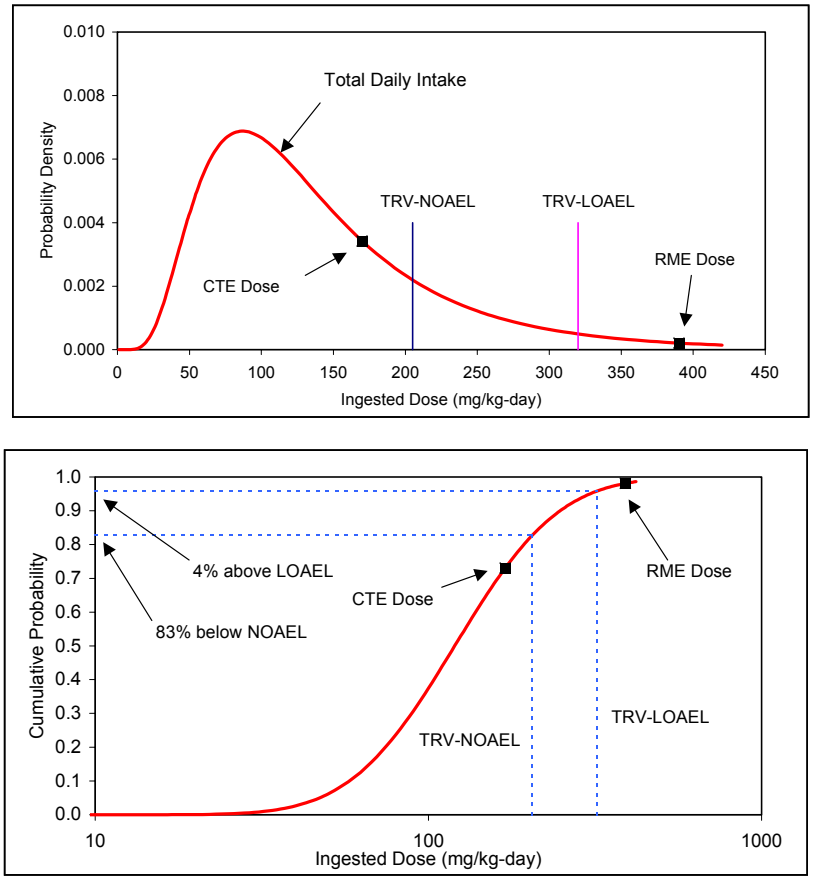


Figure 4-5. Example Comparison of Exposure Distribution to TRV.

results for a value equal to 10-times the TRV would show the fraction of the population with an HQ greater than 10).

More directly, the distribution of HQ values may be calculated by dividing each exposure value by one or all of the TRVs based on the NOAEL, LOAEL, BMDL, etc., as shown in Figure 4-6. Note that dividing a distribution by a constant does not change the shape of the distribution (only its scale), so the shape of the HQ distribution will appear identical to that of the exposure distribution. Figure 4-6 illustrates two HQ distributions; one calculated using the NOAEL-based TRV, the other using the LOAEL-based TRV. In a case such as this where there are two or more HQ distributions, a CDF format is generally easier to evaluate than a PDF format, since overlap between the curves is minimized. The CDF format allows an easy quantitative evaluation of the fraction of the population above and below any particular HQ level. For example, in the case shown in Figure 4-6, it may be seen that 83% of the population is expected to have HQ values below 1 based on the NOAEL-based TRV, while 4% are expected to have HQ values above 1 based on the LOAEL-based TRV. This type of description (percentage of the population with HQ values within a specified range) is very helpful in predicting proportions of a population exposed to specified doses of concern.

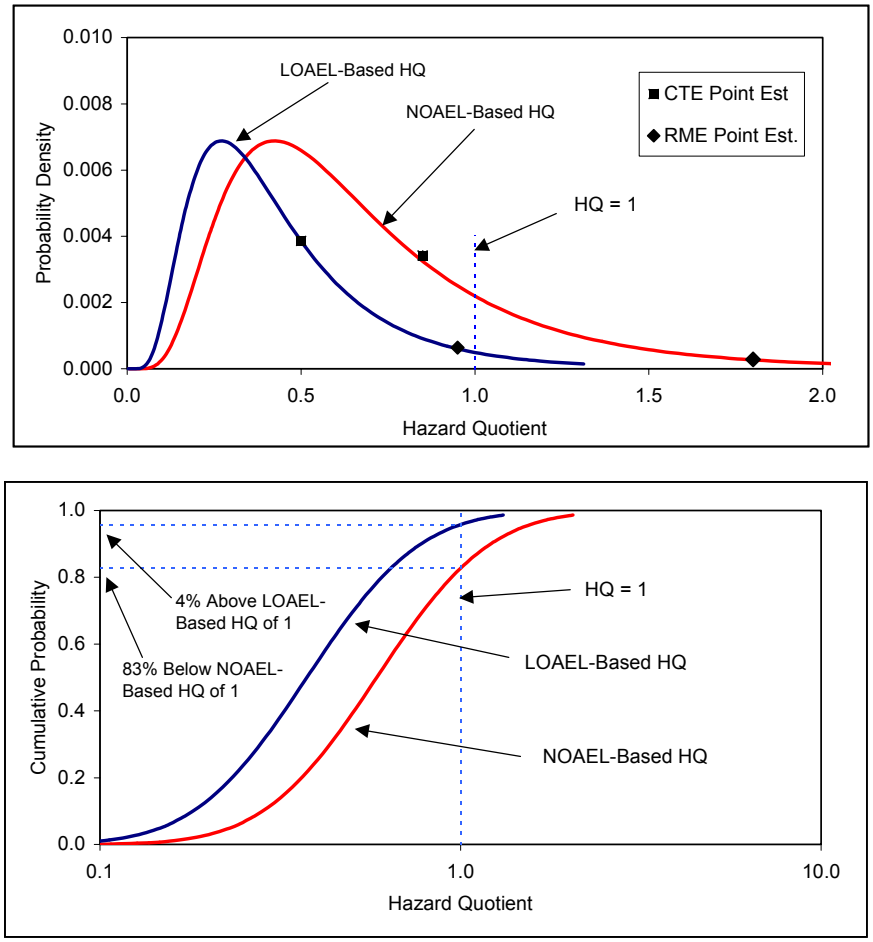
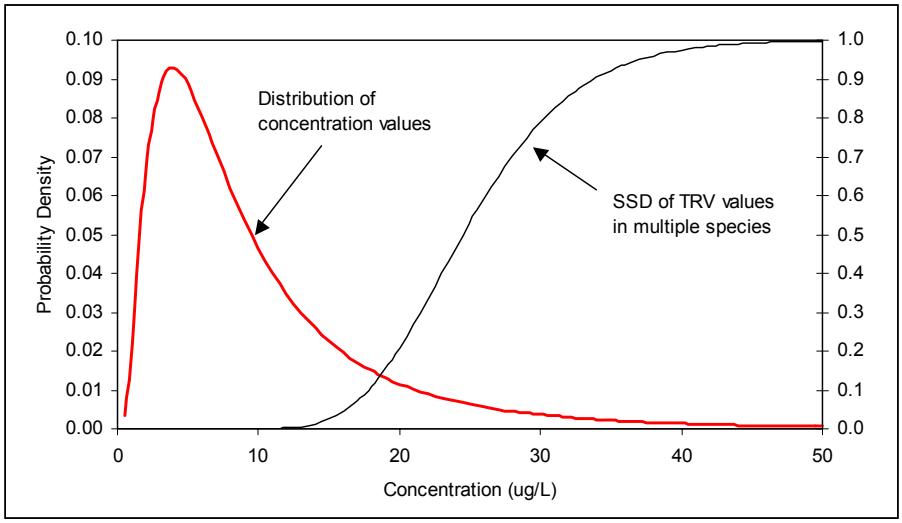
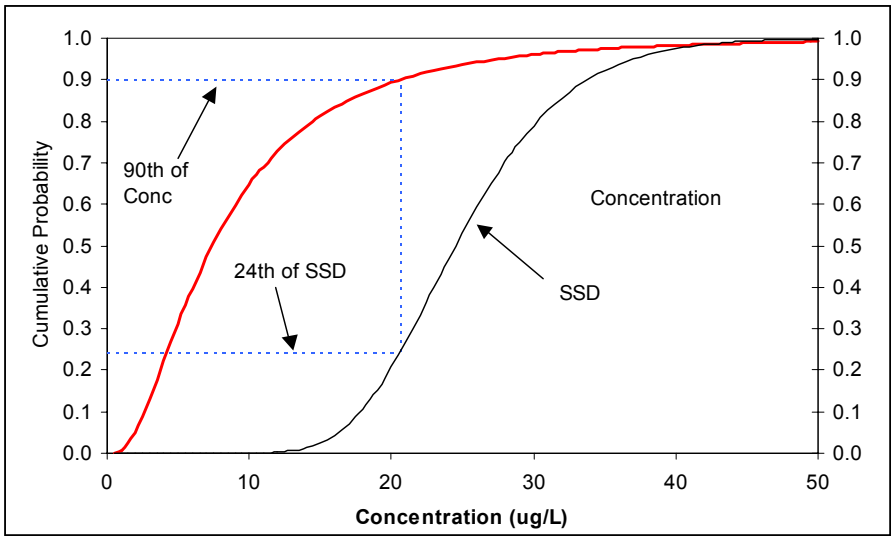


Figure 4-6. Example Distribution of HQ Values.

Variability Between Species

A similar approach may be used for characterizing the variability in risks among different species in a community. Figure 4-7 is an example that compares the distribution of concentration values in a water body (the variability might represent either time or space) to an appropriate SSD of TRVs for different species of aquatic receptors that might reside in that water body. Three different graphical formats are illustrated. In the upper panel, the PDF of concentration is compared to the CDF of the SSD. This format is easy to understand and may be interpreted visually, but is difficult to interpret quantitatively. The middle panel shows that same information, but with both distributions presented in CDF format. This allows for a quantitative evaluation of the fraction of the species that will be above their respective TRVs at any specified part of the exposure distribution. For example, using a simple graphical interpolation process (shown by the dashed lines), it may be seen that the 90th percentile of concentration (21 ug/L) will impact approximately 24% of the exposed species. The bottom panel shows the results when this same process is repeated (mathematically) for each of the concentration percentiles. As seen, hazards to the community of receptor species is quite low until concentration values reach the 80th to 85th percentile, but then rise rapidly. For example, a concentration value equal to the 95th percentile (about 28 ug/L, which will occur approximately 5% of the time) is expected to impact approximately 68% of the exposed species, and the 99th percentile (which will occur about 1% of the time) is expected to impact nearly all of the exposed species.





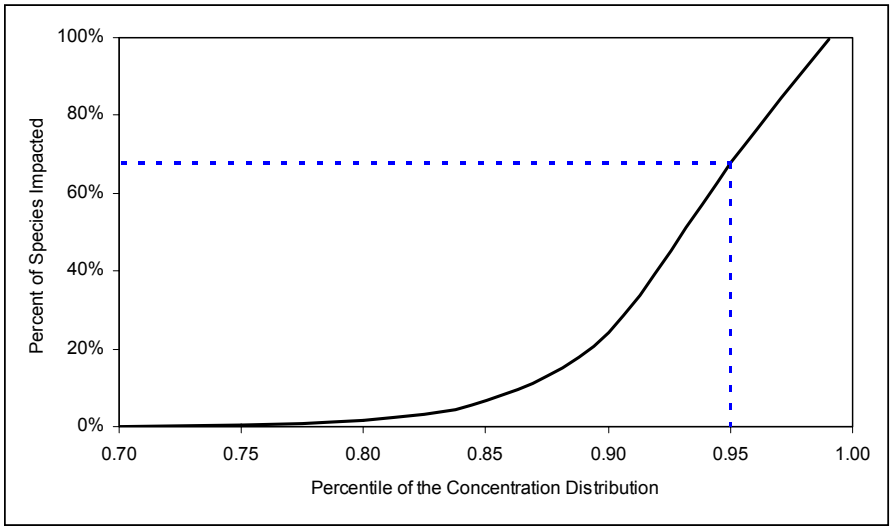


Figure 4-7. Example Presentation of Species Sensitivity Distribution.

4.6.2 VARIABILITY IN RESPONSE

As noted above, HQ and Hazard Index (HI) (where appropriate) values are a convenient way to characterize risk to ecological receptors, but interpreting the biological significance of the ranges of HQ values greater than 1 is not always easy. One of the main advantages to the PRA approach is that distributions of exposure may be combined with exposure-response distributions in order to generate distributions that characterize the frequency and magnitude (severity) of responses in an exposed population. Two examples of this approach are presented below.

Example 1: Dichotomous Response

In this hypothetical example, a toxic chemical is being transported by surface water run-off from a Superfund site into a nearby stream. Because of short-term and seasonal variability in rainfall levels (which influences both run-off rate and stream flow), the concentration of the chemical in the stream has been observed to vary as a function of time. The risk manager at the site wants to know two things: (1) How often will the concentration enter a range that can cause acute lethality in fish?; and (2) When that happens, what percent of the fish population is likely to die? Exhibit 4-6 summarizes the hypothetical concentration data and illustrates the basic approach. In this case, the concentration data are most conveniently modeled as an empirical PDF. Next, assume that the acute concentration-lethality curve is available for the chemical of interest in a relevant indicator species of fish. For convenience, assume the response function is the same as that shown in Exhibit 4-3. Then, the PDF for acute mortality may be generated by repeated sampling from the concentration distribution and calculating the probability of response (acute mortality) for each concentration value selected. Because this is a case where the entire population of fish at the exposure location may be assumed to be exposed to the same concentration in water, the probability of mortality in a single fish is equivalent to the average fraction of the population that is expected to die as a result of the exposure. The resulting PDF is shown in the graph in Exhibit 4-6. As seen, lethality is expected to be low or absent about 95% of the time, but about 5% of the time the concentration may enter a range where acute lethality may occur. The extent of mortality within the exposed population is expected to range from about 20% at the 97^{th} percentile of exposure (i.e., this is expected to occur about 3% of the time), up to about 70% at the 99^{th} percentile of exposure (i.e., this is expected to occur about 1% of the time).

EXHIBIT 4-6

MODELING VARIABILITY IN A DICHOTOMOUS RESPONSE

Scenario

Exposure of a population of fish to concentration values in a stream that vary over time

Hypothetical Concentration Data in Water

Value		Percentile
0.5	(1/2 DL)	0.00
1.1		0.10
2.5		0.25
5.1		0.50
9.2		0.75
15.8		0.90
24.7		0.95
52.6		0.99
83.1	(max)	1.00

Response

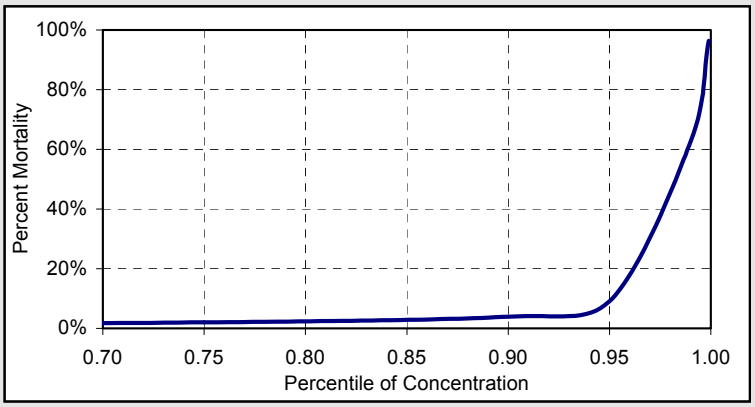
Endpoint = acute mortality
Stressor-response model fit (see Exhibit 4-2)
$P(c) = 1/(1+\exp(4.8 - 0.1*c))$

PRA Simulation

Step 1 Draw a concentration at random from the empiric distribution
Step 2 Calculate the probability of mortality at that dose
Track this as the forecast cell

Example Output

Percentile	% Lethality
0.050	0.9%
0.250	1.0%
0.500	1.4%
0.750	2.0%
0.900	3.9%
0.950	9.1%
0.990	63%
0.999	96%



Example 2: Continuous Response

Exhibit 4-7 provides a hypothetical example of modeling variability in response for a continuous endpoint. In this example, assume that a toxic chemical has been released by a Superfund site and has been transported in low levels by air to a nearby meadow. Among the receptors of potential concern in the meadow are a number of different types of small mammal, and the field mouse has been selected to serve as an indicator species for this group. The goal of the PRA is to characterize the effects of the chemical on the growth of field mice in the meadow. Exposure occurs mainly by ingestion of seeds that have been contaminated by uptake of the chemical from soil, and it has been determined that the variability in average daily intake (DI) of chemical from the diet can be modeled as a lognormal distribution with mean of 104 mg/kg-day, and a standard deviation of 127 mg/kg-day. Assume for convenience that the exposure-response curve for growth inhibition in mice by the chemical is the same as that presented previously in Exhibit 4-4. Given these inputs, the expected distribution of responses is derived as follows:

Step 1: Draw a random value for the DI of a random member of the population
Step 2: Calculate the mean response m(d) and the standard deviation of the response s(d) for a group of individuals exposed at that dose (d)
Step 3: Define the distribution of responses at that dose as NORMAL[m(d), s(d)]
Step 4: Draw a response from that distribution, and track this as the output variable

An example of the output for this example is shown in the two graphs at the bottom of Exhibit 4-7. As seen, mice that are not exposed to the chemical display a range of growth rates ranging from about +10% to +40%. Many of the mice (about 90%) residing in the contaminated field are experiencing a range of growth rates that are only slightly decreased from rates expected for unexposed animals. However, about 10% of the animals have weight gains that are markedly less than for unexposed animals, ranging from about +5% to -30% (i.e., a net weight loss of 30% compared to the starting weight).

It should be noted that the response distribution calculated in this way is what would be expected for a large population of exposed receptors. If the actual exposed population is small, then the actual response distribution may vary somewhat compared to the typical response shown in Exhibit 4-7. In cases where it is important to evaluate this variability about the expected average pattern of response, this may be done by running repeated Monte Carlo simulations using a number of trials (iterations) within each simulation that is equal to the expected size of the exposed population. Each simulation will thus represent a possible response distribution in the exposed population, and the range of responses across different populations may be evaluated by comparing the multiple simulations. As noted above, the magnitude of the variability between populations is expected to be small if the population size (number of trials) is large, although this depends on the characteristics of the exposure and response functions.

EXHIBIT 4-7

MODELING VARIABILITY IN A CONTINUOUS RESPONSE

Scenario

Exposure of a population of field mice to a chemical ingested via the food chain

Example Inputs

Exposure

Distribution of Average DI LN(104,127)

Response (see Exhibit 4-3)

Endpoint = Growth (% increase in 21 days)

Stressor-response model fit

Mean response(dose) = 23.7 - 51.4*dose^n / (228.7^5.29 + dose^5.29)

Stdev (dose) = 7.0 (constant)

PRA Simulation

Step 1 Draw a dose at random from the lognormal distribution of dose

Step 2 Calculate the mean response [m(d)] and standard deviation of the response (s(d) at that dose

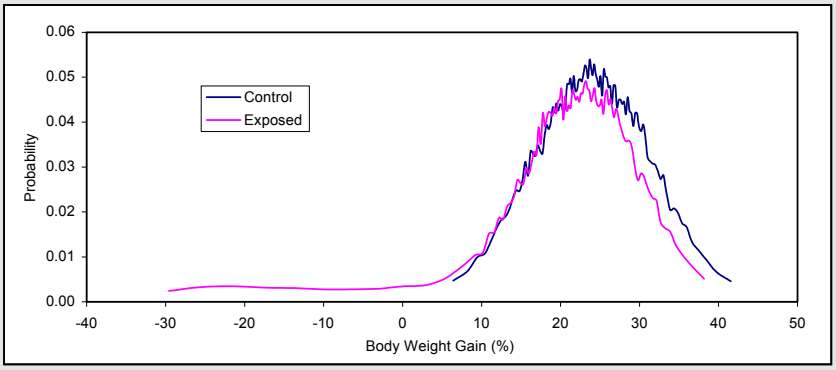
Step 3 Define the PDF for response at dose d: NORMAL(m(d), s(d))

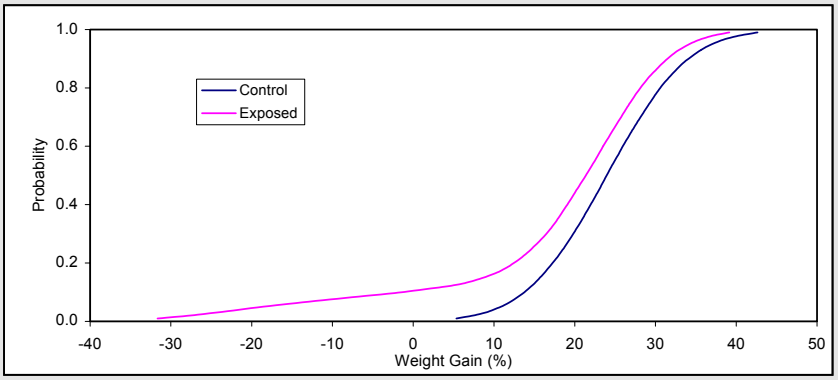
Step 4 Draw a response at random from this PDF

Track this as the forecast cell

Example Output

Percentile	Control	Exposed
0.05	10.9	-18.6
0.25	18.6	14.7
0.50	24.0	21.4
0.75	29.3	26.9
0.90	34.1	31.5
0.95	37.0	34.2
0.99	42.6	39.1





4.6.3 JOINT PROBABILITY CURVES

In this approach, if data are available to characterize the probability of a particular exposure occurring, and an exposure-response curve is available, these may be combined to yield a curve (referred to as a Joint Probability Curve) that shows the probability that a response greater than some specified magnitude will occur. An example is shown in Figure 4-8. The upper panel shows a hypothetical cumulative exposure probability distribution (plotted on the primary y-axis) along with the exposure-response curve (plotted on the secondary y-axis). The steps needed to generate the Joint Probability Curve are as follows:

Step 1: Select an exposure level "x" and record the probability (P_x) of exceeding that exposure. For example, in Figure 4-8, at an exposure of 12 units, the cumulative probability of exposure is 84%. Thus, the probability of exceeding that exposure is 16%.

Step 2: Find the expected response at that same exposure (R_x). In this case, the response at an exposure of 12 is 2.2.

Step 3: Plot a data point at R_x on the x-axis and P_x on the y-axis.

Step 4: Repeat this process for many different exposure levels, being sure to draw samples that adequately cover all parts of the probability scale.

The lower panel of Figure 4-8 shows the results obtained using the hypothetical data in the upper panel. The advantage of this format is that it gives a clear visual display of both the probability and magnitude (severity, extent) of response. Further, the area to the left of the curve is a relative index of the population-level or community-level risk, and comparison of this area across different scenarios is helpful in comparing different risk scenarios (both in risk characterization and risk management). However, this approach is based on the mean response at a dose, and does not account for variability in response between multiple individuals all exposed at that dose. Employing a two-dimensional Monte Carlo analysis (2-D MCA) procedure could help to display this variability in response between the individuals at a given dose.

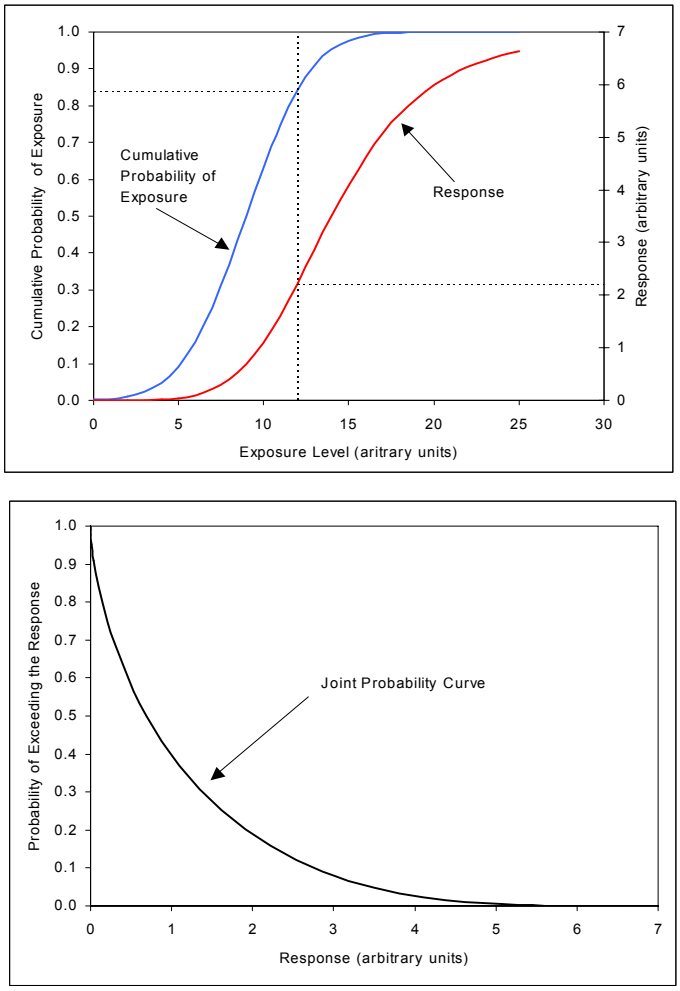


Figure 4-8. Example Joint Probability Curve.

Note that unless 2-D MCA is used, this approach does not require Monte Carlo modeling. Rather, the calculations can usually be performed in a spreadsheet format using built-in spreadsheet functions.

4.7 MODELING UNCERTAINTY IN ECOLOGICAL RISK ASSESSMENTS

As emphasized above, one of the greatest potential benefits of the PRA approach is the ability to combine estimates of uncertainty associated with different components of the exposure and risk models in order to describe the overall uncertainty in final exposure or risk estimates. Some basic options for characterizing and presenting uncertainty in exposure, toxicity, HQ, and response are presented below.

4.7.1 UNCERTAINTY IN EXPOSURE

Most estimates of dose-based exposure for terrestrial receptors (birds, mammals) are based on calculated estimates of chemical intake using simple or complex food web models, sometimes coupled with environmental fate and transport models that can link risk to a receptor with a source of contamination. In cases where receptors are exposed mainly by direct contact rather than ingestion (e.g., fish, soil invertebrates, etc.), concentration-based (as opposed to dose-based) descriptors of exposures may be derived using mathematical fate and transport models. The basic principles for modeling uncertainty in ecological exposure models (either dose-based or concentration-based) are the same as discussed in Appendix D. In brief, probability distribution functions of uncertainty (PDFu's) are used to characterize the uncertainty in the parameters of the probability distribution functions of variability (PDFv's) for some or all variables in the exposure model. Then, a 2-D MCA is used to derive quantitative estimates of the uncertainty around each percentile of the variability distribution of exposure. Figure 4-9 illustrates the type of tabular and graphic outputs that this approach generates.

Variability Percentile	Uncertainty Percentiles		
	5th	Mean	95th
0.05	0.4	1.1	2.0
0.10	0.7	1.6	2.8
0.15	0.9	2.1	3.5
0.20	1.2	2.6	4.2
0.25	1.5	3.1	5.0
0.30	1.8	3.7	5.9
0.35	2.1	4.3	6.7
0.40	2.6	5.0	7.6
0.45	3.0	5.8	8.7
0.50	3.6	6.6	9.9
0.55	4.2	7.7	11.3
0.60	5.0	8.8	12.9
0.65	5.9	10.3	14.8
0.70	7.2	12.1	17.2
0.75	8.8	14.4	20.3
0.80	10.9	17.5	24.1
0.85	14.5	22.0	30.1
0.90	20.1	29.6	39.4
0.95	32.9	46.5	60.0

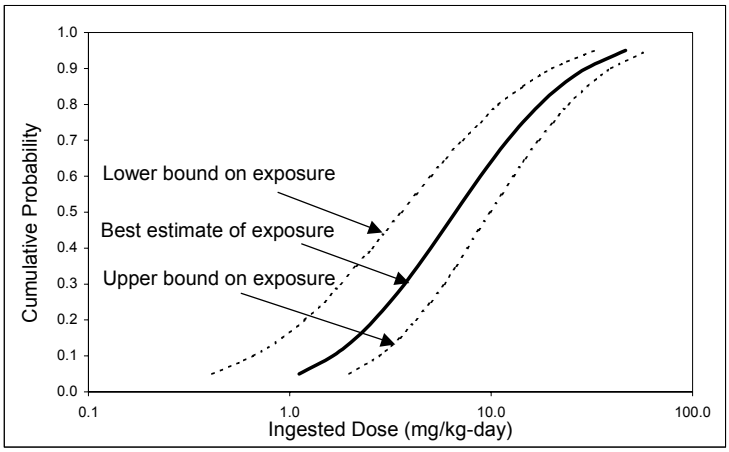


Figure 4-9. Example Presentation of Uncertainty in Exposure.

If exposure is based on measured rather than calculated values by PRA (e.g., measured concentrations in an abiotic medium,

measured concentrations in receptor tissues), uncertainty in the empirical or best-fit continuous distribution through the data can be quantified using the statistical methods detailed in Appendix B.

As discussed in Chapter 1, it is important to understand that there are many sources of uncertainty and that this approach to uncertainty analysis focuses mainly on parameter uncertainty and uncertainty in the true shape of input variable distributions. It does not capture other sources of uncertainty relating to the fundamental adequacy of the exposure and risk models used to describe the behavior of complex biological systems or of sampling and analytical errors and uncertainties, so the uncertainty estimates should always be interpreted in this light as being somewhat incomplete.

4.7.2 UNCERTAINTY IN TOXICITY

Toxicity information used for ERAs is often a source of uncertainty in the risk assessment process. This uncertainty may arise from multiple areas and may include both quantitative uncertainty in the dose-response data (involving toxicokinetics and study designs) and qualitative uncertainty in the relevance of the data (involving toxicodynamics). Methods for characterizing the quantitative uncertainty in both point estimates of toxicity (TRVs) and in full exposure-response curves are outlined below.

Uncertainty in TRVs

TRVs for a chemical are point estimates of exposure levels that do not cause an unacceptable effect in an exposed receptor population. Ideally, all TRVs would be based on NOAEL and LOAEL values derived from studies in which the receptor, endpoint, exposure route and duration were all matched to the assessment endpoints defined for the site. However, such exact matches are seldom available. Therefore, it is often necessary to extrapolate available toxicity data across route, duration, endpoint and/or species, leading to uncertainty in the most appropriate value to use as the NOAEL or LOAEL. There are no default methods for developing TRVs on a site. However, some options include the use of allometric dose scaling models, physiologically-based biokinetic models, benchmark dose estimates or other approaches based mainly on policy and/or professional judgment. Guidelines for dealing with the uncertainty in components of the TRV derivation by uses of PRA are provided below.

Uncertainty in NOAELs and LOAELs

Uncertainty in the NOAEL or LOAEL for a chemical has two components: (1) uncertainty within a study; and (2) uncertainty between studies, under exact specified conditions of exposure.

Assuming that a single study has been selected to provide the NOAEL and/or LOAEL values to be used in deriving a TRV for a chemical, it is customary to define the NOAEL as the highest exposure that did not cause a statistically significant effect, and the LOAEL is the lowest exposure that did cause a statistically significant effect. As noted earlier (see Section 4.5.1), this approach has a number of limitations, and there may be substantial uncertainty as to whether the observed NOAEL and LOAEL values actually bracket the true threshold effect level. One way to quantify uncertainty in the exposure levels that cause some specified level of adverse effect is through the use of exposure-response curve-fitting software such as EPA's BMDS package. In this approach, the risk assessor selects some level of effect that is judged to be below a level of concern, and another level of effect that would be of concern. The choice of these response levels is a matter of judgment, and depends on the nature and severity of the endpoint being evaluated. A specified level of effect is referred to as a Benchmark Response (BMR), and the exposure that causes that response is referred to as the Benchmark Dose (BMD). Given information on the number of test organisms in each test group and on the variability of the response in those

organisms, the BMD software uses maximum likelihood methods to derive the 5% lower confidence bound on the exposure that causes the BMR. This is referred to as the BMDL. This uncertainty bound may be used to quantify the uncertainty in the BMD, and hence to characterize this source of uncertainty in the TRV. The simplest method for approximating the uncertainty distribution around the BMD is to assume the distribution is approximately normal, with mean equal to the BMD and standard deviation (standard error) given by:

Stdev=(BMD - BMDL) / 1.645

For advanced analyses, a more accurate characterization of the uncertainty distribution around the BMD may be derived by Monte Carlo simulation. In this approach, each model parameter is assumed to be normally distributed, with mean and standard error values provided by the BMDS output. Monte Carlo simulation is then used to select alternative model parameter sets, being sure to account for the covariance between parameters (the covariance matrix is also provided by the BMDS output). For each parameter data set, the BMD is calculated, and the distribution of BMD values across many iterations is a better approximation of the uncertainty in the BMD.

Uncertainty in the effect level (NOAEL or LOAEL) for a chemical may also arise because there is more than one study available for the chemical, and the studies do not yield equal estimates of the effect level. It is important to note that the process of reviewing available toxicity studies, choosing the most relevant endpoint for use in deriving a TRV, and identifying the most relevant study is a process requiring basic toxicological expertise (not probability or statistics), and this process must be completed both for point estimate and probabilistic risk assessments. In general, studies based on different receptors, endpoints, exposure routes and/or durations are not equally relevant for evaluating a particular assessment endpoint in a particular indicator species. However, in some cases, multiple studies of the same endpoint in the same species will be available. In such a case, assuming that all the studies are judged to be equally reliable, the best estimate of the LC50 may be derived by calculating the geometric mean of the available alternative values (after adjustment to constant hardness). Uncertainty around the best estimate may then be based on the observed inter-study variability, using the basic principles for choosing PDFu's as described in Appendix B.

Uncertainty in Extrapolation of TRVs

In general, extrapolation of TRVs across species or endpoints is not desirable, since the magnitude and direction of any potential error is generally not known. Sometimes, extrapolations between species are attempted based on allometric scaling models that seek to adjust toxicity values accounting for differences in body weight. Alternatively, physiologically-based pharmacokinetic (PBPK) models that seek to account for differences in a number of other physiological variables (metabolism rate, organ size, blood flow, etc.) can be used. However, the validity of these models is often not well established. In those cases where these models are used, and where the uncertainty in the model is judged to warrant quantitative evaluation, the primary source of the model should be consulted in order to derive an estimate of the uncertainty in the quality of the extrapolation and in the parameters of the model. As noted earlier, PRA may capture uncertainty associated with model input parameters, but does not usually capture all sources of uncertainty in the model. In particular, most models of this sort are designed to extrapolate only the average response as a function of dose, and are not intended to extrapolate variability between individuals at a specified dose. When no mathematical model is available to support quantitative extrapolation across species, exposure duration or endpoint, professional judgment and/or policy may be used to select extrapolation factors to account for the uncertainty.

The risk assessor should ensure that the risk manager understands the uncertainty associated with any model selected and applied, and that the results of the calculations (point estimate or PRA) are conditional upon the model selected.

Uncertainty in Parameters of the Dose-Response Models

When toxicological exposure-response data are fit to mathematical equations, the fitting software will usually provide quantitative information on the uncertainty in the best estimates for each of the model parameters. For example, in the dichotomous model illustrated in Exhibit 4-3, the output from the BMDS software included the following information on the uncertainty in the parameters of the best-fit logistic equation:

Parameter	Best Est	Std Error (SE)
a	-4.80	0.83
b	0.101	0.019

Because the uncertainty in the best estimate of each model parameter is asymptotically normally, uncertainty in the parameters may be modeled as:

PDFu (parameter i)=NORMAL(best estimate of parameter i, SE of parameter i)

Note that the parameters of the model are generally not independent, and generally should not be treated as such. Thus, when modeling the uncertainty in the parameters of the best-fit exposure-response model, the PDFv's for the parameters should be correlated according to the correlation matrix or the variance-covariance matrix, as provided by the modeling software.

4.7.4 UNCERTAINTY IN RESPONSE

If the risk characterization phase of the risk assessment focuses on an estimation of the distribution of responses rather than the distribution of HQ values, the uncertainty in the distribution of responses can be evaluated using two-dimensional Monte Carlo techniques using PDFu's for the parameters of the exposure and exposure-response models derived as described above. The same graphical output may be used for this presentation as was illustrated in Figure 4-9, except that the x-axis is response rather than HQ. This format is illustrated in Figure 4-10 for a dichotomous endpoint (e.g., acute lethality). In this example, the average probability of response among the members of the exposed population (shown in the graph by the black diamond symbols) is 8.2%, with a confidence bound around the mean of 4.9 to 12.8%. This is equivalent to concluding that about 8.2% of the population is expected to suffer acute lethality, but the true fraction dying could range from 4.9 to 12.8%.

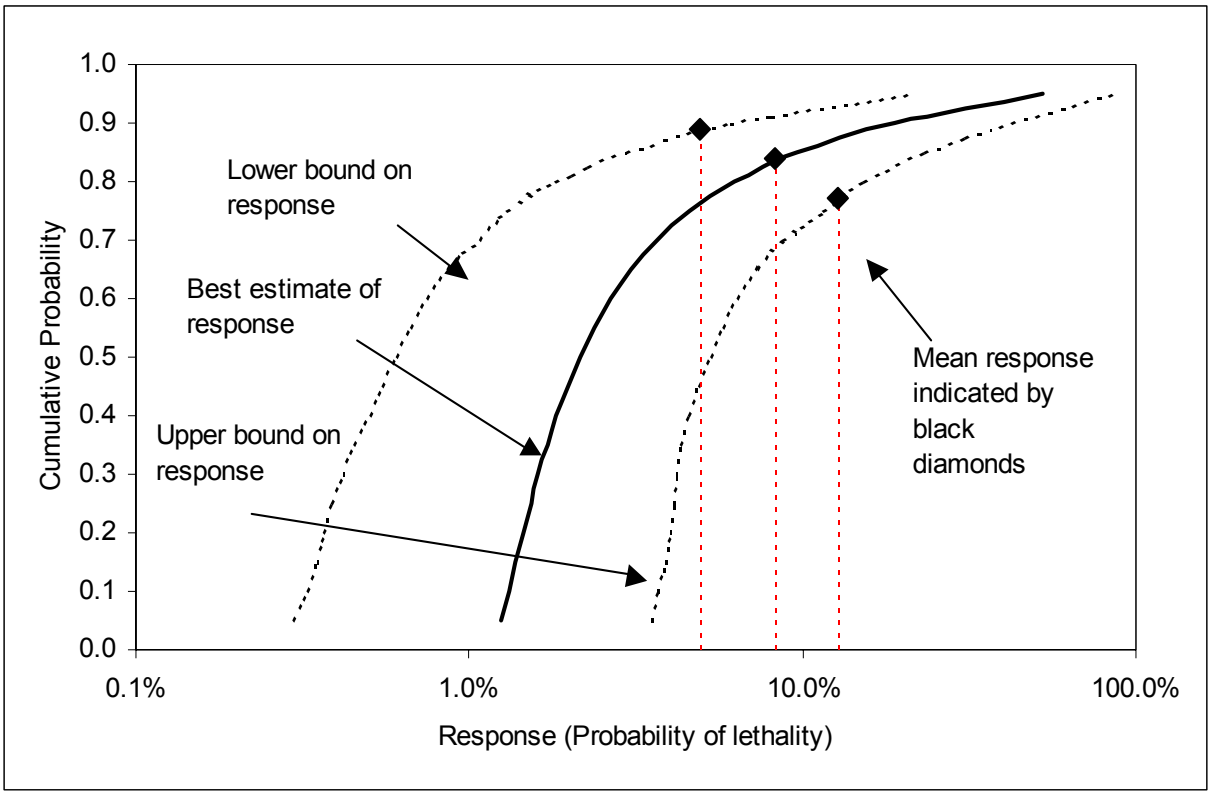
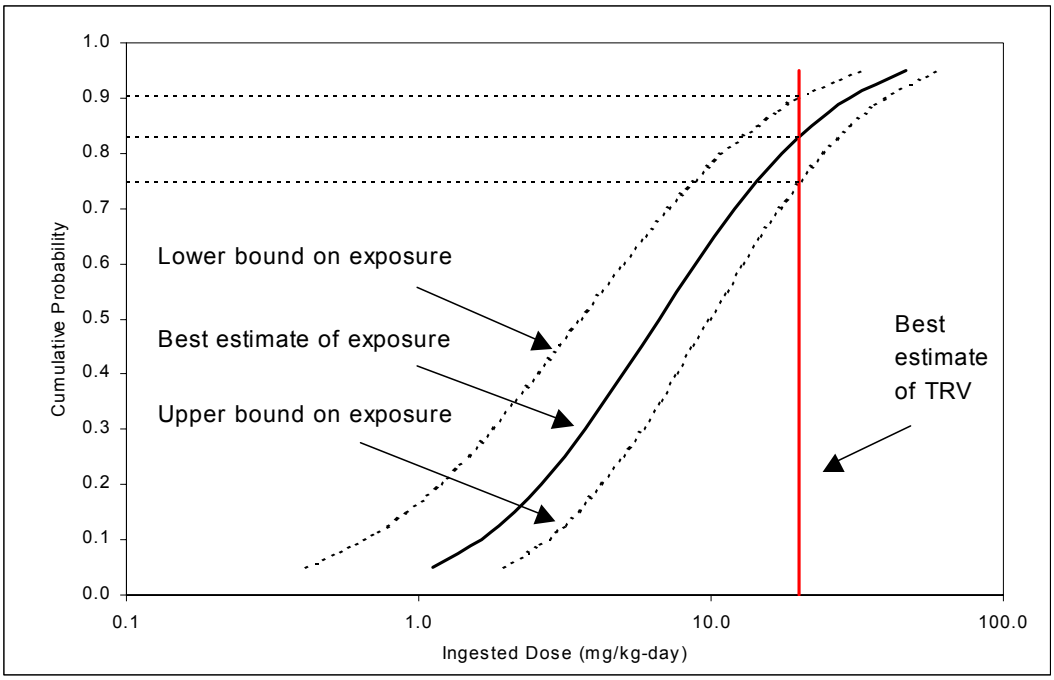


Figure 4-10. Example Presentation of Uncertainty in Response.

4.7.3 UNCERTAINTY IN HAZARD QUOTIENT

Once the uncertainty in exposure and/or toxicity distributions has been characterized as described above, there are a number of options for presenting the resultant uncertainty in the HQ (or HI, if appropriate and applicable for summing HQs) distributions. Figure 4-11 shows one simple graphical format, where the point estimate of the TRV is superimposed on the uncertainty bounds of the exposure distribution (upper panel), or the uncertainty bounds of the TRV are superimposed on the best estimate of exposure (lower panel). One could also superimpose the range of TRVs over the range of exposures, to capture most of the uncertainty in the HQ. Furthermore, such distributional outputs should always show the point estimate ranges of CTE and RME exposures in respect to the ranges of TRVs, for use in weight-of-evidence to help interpret the PRA and point estimate results. The advantage of this format is that no additional Monte Carlo modeling is needed to derive initial descriptors of uncertainty in risk. For example, in the upper panel it may be seen that the best estimate of the fraction of the population exposed at a level below the TRV is about 83%, but that this is uncertain due to uncertainty in the exposure estimates, and the true percent below the TRV might range from 74 to 90%. Similarly, in the bottom panel, the best estimate of the fraction of the population below the TRV is also about 83%, but due to uncertainty in the TRV the actual value could range from 64 to 91%. Uncertainty could also be presented by showing a combined graph with both ranges of exposure and TRVs, such as described below.



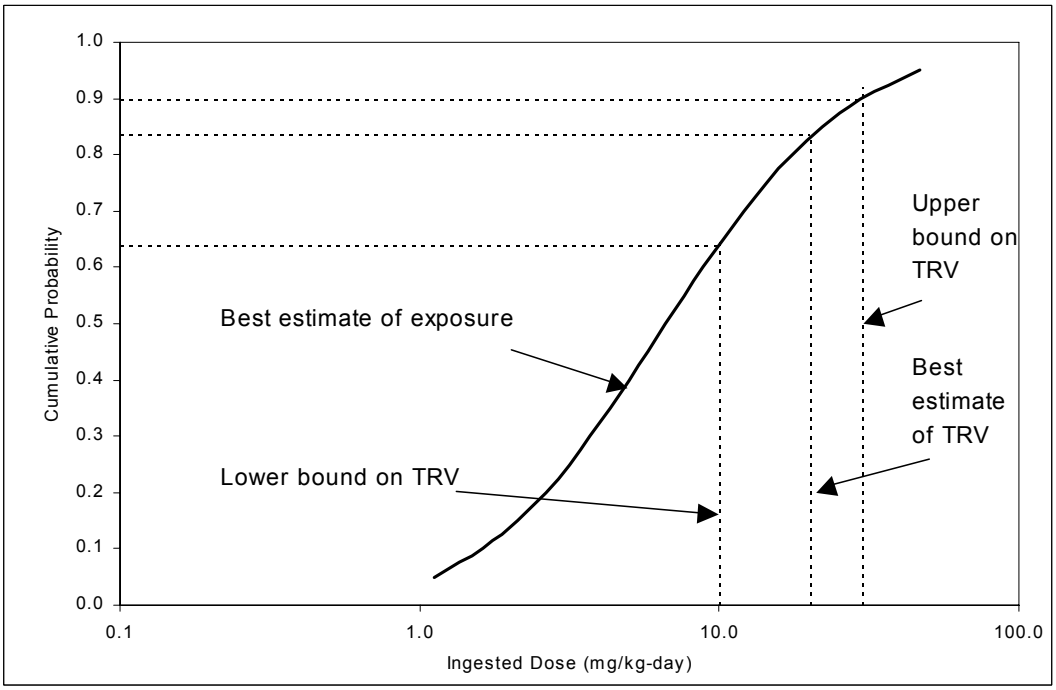


Figure 4-11. Example Presentation of Uncertainty in Exposure and TRV.

A more complete characterization of uncertainty in HQ may be achieved by using PRA to combine the uncertainty in both the exposure and the TRV terms, resulting in the uncertainty bounds on the HQ distribution itself (see Figure 4-12). In this example, it may be seen that 63% of the exposed population is estimated to have an HQ below 1.0, but that this is uncertain due to uncertainty in both the exposure distribution and the TRV, and that the true fraction of the population below a level of concern (HQ < 1) could range from 45 to 81%.

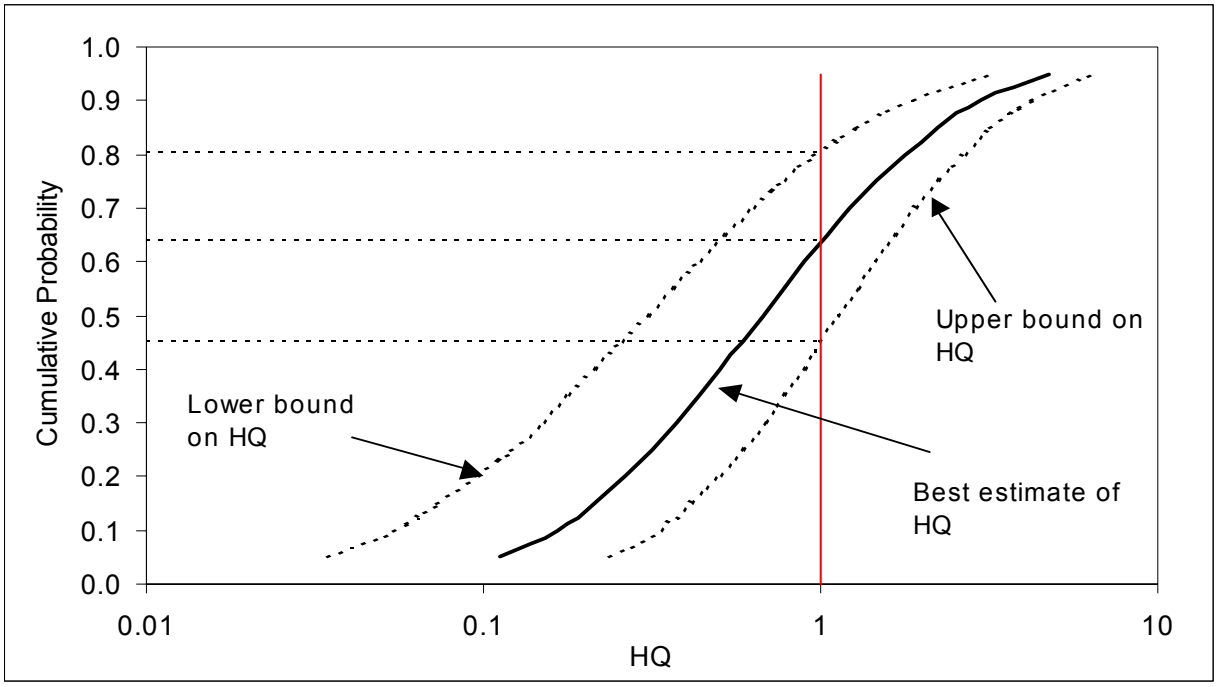


Figure 4-12. Example Presentation of Uncertainty in HQ Estimates.

4.8 INTERPRETING RESULTS OF AN ECOLOGICAL PRA

In some cases, the information contributed by a PRA may provide a more complete characterization of risks to a population of receptors than can be obtained by using point estimate methods. However, whether by PRA or by point estimate or a combination, the results of the risk assessment must be interpreted to reach a risk management decision.

In contrast to the case for human health risk assessments (where default risk-based decision rules are well established), there are no established default decision rules for identifying when risks to ecological receptors are and are not of concern. In the point estimate approach, EPA guidance (U.S. EPA 1992b, 1995) recommends consideration of both the RME and CTE exposure/dose estimates along with TRVs based on both LOAELs and NOAELs (U.S. EPA 1997a) to reach a risk management decision. The same principle applies to probabilistic ERAs.

In some cases, interpretation of an ecological PRA is relatively simple. For example, if the distribution of HQ values calculated using an appropriate NOAEL-based TRV are less than 1.0 for nearly all members of the population, then it is likely that risks are within an acceptable range for the population. Conversely, if the distribution of HQ values calculated using a LOAEL-based TRV are significantly greater than 1.0 for most members of an exposed population, then it is likely that risks are not acceptable

for the population. However, for cases which fall between these bounding conditions (and for cases where one needs to clearly define the boundaries of potential excess risks for a gradient of contamination and exposures), the level of risk or response that is considered acceptable must be defined by the risk assessor and the risk manager on a site-specific and receptor-specific basis. This evaluation should take the following factors into account:

(1) The Risk Management Goal

The risk management objective for most Superfund ERAs is defined as population sustainability (U.S. EPA, 1999). In this case, harm to some members of the exposed population may be acceptable, if that harm does not lead to an overall reduction in population viability. This situation (protection of a population rather than protection of individuals) is sometimes equated with use of the CTE (average) receptor as the basis for risk management decision making. That is, if the HQ for the CTE receptor is below a level of concern, it is sometimes assumed that population risks are acceptable.

However, the choice of the CTE receptor as the basis for risk management decision making may not be sufficiently protective in all cases. For the vast majority of wild populations, the proportion of the population that must be protected to ensure population stability will be unknown. At a small number of sites, a population biologist may be able to provide some information. Moreover, the percentile of the CTE receptor in the exposure or risk distribution may vary depending on the shape of the distribution. The proportion of the population experiencing exposure greater than that of the CTE receptor could range from less than 10% up to 50% or even higher. Also, the ecological significance of an adverse effect on some members of a population depends on the nature of the stressors and on the life history and population biology of the receptor species. Because of these complexities, use of the CTE as a decision threshold for nonthreatened or endangered species may be appropriate in a small number of cases, but risk assessors and risk managers should realize that the choice of the CTE receptor requires a species- and endpoint-specific justification and the CTE should not be used as the default basis for a risk management decision. Rather, for the majority of ERAs, the risk management decision should be based on the RME receptor or an upper percentile of the distribution of variability in risk/exposure.

(2) The Toxicological Basis of the TRV

The biological significance of a distribution of variability in HQ cannot be interpreted without a proper understanding of the nature of the TRV being used to evaluate the distribution. This includes the nature of the toxicological endpoint underlying the TRV, its relevance to the assessment endpoint, and the shape (steepness) of the dose-response curve. For example, an HQ of 2 based on an EC20 for reduction in reproductive success would likely be interpreted as more significant toxicologically than an HQ of 2 based on the EC20 for an increase in liver weight. Likewise, an HQ of 2 based on an LC_{low} for acute lethality would be more significant if the dose-response curve for lethality were steep than if it were shallow, since it would be easier to cause greater response with smaller increases in exposure to contaminants.

(3) The Characteristics of the Receptor

Ultimately, the question which must be assessed is whether an effect of degree "x" occurring in "y" percent of the population is biologically and ecologically significant. This, in turn depends on the attributes of the receptor being evaluated. For example, a reduction of 10% in the reproductive success of a fecund and common species (e.g., the field mouse) might not lead to a significant reduction in population number, while the same effect could be of concern in a species with lower fecundity and/or lower population density (e.g., the moose). Thus, the interpretation of an analysis of variability in exposure and/or effect often requires the input of a trained population biologist with expertise in the receptor of concern.

Because of these issues, there is no default rule for what level of effect is and is not acceptable for an exposed ecological population; except for the case of no potential excess risks where the RME exposures do not exceed the TRV based on a NOAEL, assuming there is reasonable confidence in those exposure and toxicity values. In some cases, mathematical models may be available for predicting the population-level consequences of a given pattern of effects (e.g., see ECOFRAM 1999a for some aquatic population models), but in general the extrapolation from a distribution of individual responses to an estimation of population-level effects is difficult. For this reason, close consultation between the risk manager and the ecological risk assessor is necessary for translating results of an ERA into an appropriate and successful risk management decision.

4.9 GUIDELINES FOR PLANNING AND PERFORMING A PROBABILISTIC ERA

4.9.1 PLANNING AN ECOLOGICAL PRA

Chapter 2 provides a general discussion of the key steps that should be followed when planning a PRA. These guidelines are equally applicable to ecological PRA as to human health PRA. Of the key steps in the process, most important are the following:

Dialogue Among Stakeholders

As discussed in Section 4.2, the decision if and when to perform an ecological PRA is an SMDP shared by risk assessors, risk managers, and stakeholders, including members of the public, representatives from state or county environmental agencies, tribal government representatives, natural resource trustees, private contractors, and potentially responsible parties (PRPs) and their representatives. A scoping meeting should be held after the completion of the baseline risk assessment in order to discuss the potential purpose and objectives of a PRA, and to identify the potential value of the analysis to the risk management process. If it is decided to perform at least an initial PRA evaluation, subsequent meetings of a similar type should occur iteratively in order to assess whether any further effort is warranted.

Preparation of a Workplan

Any PRA beyond the simplest screening level evaluation should always be accompanied by a workplan. The purpose of the workplan is to ensure that all parties agree on the purpose and scope of the effort, and on the specific methods, data, and procedures that will be used in the PRA. Workplans should be developed according to available guidance for workplans for nonprobabilistic ERA (U.S. EPA, 1992b,

1997a) and should consider three elements: (1) the 16 guiding principles of MCA (U.S. EPA, 1997b); (2) the eight guidelines for PRA report submission (U.S. EPA, 1997b); and (3) the tiered approach to ERA (U.S. EPA, 1997a). Development of a workplan for PRA is discussed in greater detail in Chapter 2, and Exhibit 4-8 summarizes the key elements of a proper workplan. The workplan must be submitted to the BTAG coordinator and/or regional ecotoxicologist for review and for approval by the risk manager. The EPA strongly recommends that PRPs who wish to perform PRAs of ecological risk involve the Agency in the development of a workplan in order to minimize chances of significant disagreement, as is required by EPA policy.

EXHIBIT 4-8

EXAMPLE ELEMENTS OF A WORKPLAN FOR ECOLOGICAL PRA

1. Introduction/Overview
   - Conceptual site model
   - Assessment endpoints
   - Indicator species
   - Measures of exposure and effect

2. Description of Exposure and Risk Models
   - Basic exposure models (fate and transport, uptake, food web, intake, etc.)
   - Basic risk models (HQ, dichotomous response, continuous response)

3. Results from a Point Estimate Assessment
   - CTE and RME risk estimates from baseline evaluation

4. Rationale why a PRA will be helpful
   - Goals of the assessment (variability, uncertainty, both)
   - Expected benefit to risk manager

5. Description of the Proposed PRA
   - Exposure scenarios to be evaluated
   - Output variables to be modeled in variability and/or uncertainty space

6. Proposed PDFs, and their basis
   - Method for performing sensitivity analysis and for selecting key variables
   - Data source for characterizing key variables
   - Approach for selecting and parameterizing key variables
   - Proposed list of PDFs for exposure variables (optional but desirable)
   - Method for dealing with the concentration term
   - Method for dealing with correlations

7. Proposed Software and Simulation Approach
   - Commercial or custom
   - Monte Carlo or Latin Hypercube
   - Number of Iterations
   - Method(s) for sensitivity analysis

8. Preliminary Results (optional, but helpful)
   - Results of a screening level evaluation
   - Identification of variables where more effort is needed to improve the distribution function

4.9.2 EVALUATING AN ECOLOGICAL PRA

When an ecological PRA is submitted to EPA for consideration, it will be reviewed in order to determine if it has been performed in accord with sound principles of ERA (U.S. EPA, 1997a, 1998), and with sound principles of PRA (U.S. EPA, 1997b). A general checklist that may be helpful to reviewers is provided in Appendix F, and key features of this checklist are summarized in Exhibit 4-9. Eight specific conditions for acceptance of a PRA submitted to EPA are provided in U.S. EPA (1997b).

At the discretion of EPA risk assessor or risk manager, the PRA report may be submitted for additional EPA internal review and/or an external review process in accord with Agency guidelines for conducting peer reviews (U.S. EPA, 2001). The external peer review may be used in cases where the issues are complex or contentious and the opinions of outside expert peer reviewers can improve the PRA.

EXHIBIT 4-9

CHECKLIST FOR INCLUDING A PRA AS PART OF THE ERA (SEE APPENDIX F)

- All risk assessments should include point estimates prepared according to current Superfund national and regional guidance.
- A workplan must be submitted for review and approval by the appropriate EPA regional project manager (RPM) and/or BTAG coordinator prior to submission of the PRA.
- A tiered approach should be used to determine the level of complexity appropriate for the ERA. The decision to ascend to a higher level of complexity should be made with the risk manager, regional risk assessor and other stakeholders.
- The eight conditions for acceptance presented in the EPA policy on PRA (U.S. EPA, 1997b) should be clearly addressed by each PRA submitted to the Agency.
- Information in the PRA should possess sufficient detail that a reviewer can recreate both the input distributions and all facets of the analysis. This includes copies of published papers, electronic versions of necessary data and other materials deemed appropriate by EPA.

4.10 EXAMPLE OF THE TIERED PROCESS IN ERA

As discussed in detail in Chapter 2, one of the key elements in the risk assessment process is deciding if and when further analysis is warranted. This includes decisions regarding whether to employ PRA calculations to supplement point estimate calculation, and if so, what level of effort to invest in those PRA calculations. The following section presents a relatively simple hypothetical example illustrating how the tiered approach might operate at a site where ecological risk is an important concern.

Problem Formulation

PestCorp is a former chemical manufacturing facility that produced mainly chlorinated pesticides 10 to 20 years ago. Data collected on the PestCorp property indicate that a number of spills or releases of chlorinated pesticides took place when the facility was in operation, and that site soils are broadly contaminated, especially with pesticide X. This contaminated soil has lead to impacts on a nearby lake of about 300 acres that receives surface water runoff from the PestCorp site. Samples from the lake reveal low but detectable levels of pesticide X in water, with relatively high values in sediment and in the tissues of a variety of aquatic organisms (crayfish, snails, benthic macroinvertebrates and fish). The concentration values in all media (water, sediment, aquatic organisms) tend to be highest in the part of the lake receiving runoff from the PestCorp property, with a gradient of diminishing values at locations further away from the area where runoff enters the lake.

A BTAG committee formed by EPA to identify potential ecological concerns at the site recognized that many different species could be exposed to the contaminants in the lake, including aquatic receptors residing in the lake (fish, invertebrates, aquatic plants), as well as mammals and birds that frequent the lake for food or water. Because pesticide X is lipophilic and tends to biomagnify in the food web, the BTAG decided that the highest risks would likely occur in higher-level predators such as mammalian omnivores, and selected the racoon as a good indicator species to represent this trophic group. Pathways of exposure that were identified as warranting quantitative evaluation included (a) ingestion of water, (b) ingestion of aquatic food items, and (c) incidental ingestion of sediment while feeding or drinking at the lake. The BTAG determined that the assessment endpoint was protection of mammalian omnivore populations.

Point Estimate Risk Evaluation

A series of iterative screening-level point estimate calculations (Steps 1 to 2 of the 8-step ERAGS process) were performed to investigate whether or not there was a basis for concern at the site. Initial calculations using simplified and conservative inputs (i.e., exposure based on the maximum measured concentration in each medium, an area use factor of 1, and the most conservative available TRVs) indicated that the HQ value for pesticide X could be quite large. Therefore, a refined screening level evaluation was performed in which point estimates of CTE and RME risk were derived using the best information currently available. Key elements of the approach are summarized below:

- The CTE receptor was assumed to be exposed at a location where concentration values were the average for the whole lake, and the RME receptor was assumed to be exposed at a location where concentrations were equal to the 95^{th} percentile of values from the lake.

- Because only limited data were available for measured concentrations of pesticide X in aquatic prey items, the concentration values in aquatic prey were estimated using a linear bioaccumulation model: C(prey)=C(sed) x BAF. The BAF was estimated from the existing data by finding the best fit correlation between the concentration values in sediment and crayfish at 7 locations in the lake: C(crayfish)=5.04 x C(sed) (R^2=0.792).

- The TRV values were based on a study in mink in which the toxicity endpoint was the percent inhibition of reproductive success.

These inputs and the resulting HQ values are shown in Exhibit 4-10. As seen, estimated risks to the CTE receptor approach or slightly exceed a level of concern (HQ=4.7E-01 to 1.4E+00), and risks to an RME receptor are well above a level of concern (9.1E+00 to 2.7E+01). The chief pathway contributing to the dose and risk is ingestion of contaminant in aquatic food web items (crayfish, fish, amphibians, etc.).

EXHIBIT 4-10

REFINED SCREENING POINT ESTIMATE INPUTS AND RESULTS

Basic model

HQ = DI(total) / TRV
DI(total) = DI(water) + DI(food) + DI(sed)
DI(i) = C(i) * IR(i)*AUF(i)

Other Assumptions

C(diet) = C(sed) * BAF
IR(sed) = IR(diet) * F(sed)
IR(diet) = IR(total) * F(diet)

Category	Variable	Variable	Units	Point Est. Values	
				CTE	RME
Inputs	Concentration	Concentration in water	mg/L	0.12	0.38
		Concentration in sediment	mg/kg	24	77
		BAF (sediment to aquatic prey)	--	5	5
		Concentration in aquatic prey	mg/kg	120	385
	Intake Rates	Total water intake rate	L/kg-day	0.082	0.12
		Total food intake rate	kg/kg-day	0.06	0.09
		Fraction of diet that is sed	--	0.03	0.06
		Fraction of diet that is aquatic prey	--	0.15	0.25
	Area Use Factors	Fraction of total water ingested at the lake	--	0.3	0.6
		Fraction of total diet from the lake	--	0.25	0.6
	TRVs	LOAEL-based TRV	mg/kg-day	0.6	0.6
		NOAEL-based TRV	mg/kg-day	0.2	0.2
Results	Daily Intake	Water ingestion	mg/kg-day	3.0E-03	2.7E-02
		Sediment ingestion	mg/kg-day	1.1E-02	2.5E-01
		Aquatic prey ingestion	mg/kg-day	2.7E-01	5.2E+00
		Total	mg/kg-day	2.8E-01	5.5E+00
	HQ (LOAEL-Based)	Water ingestion		4.9E-03	4.6E-02
		Sediment ingestion		1.8E-02	4.2E-01
		Aquatic prey ingestion		4.5E-01	8.7E+00
		Total		**4.7E-01**	**9.1E+00**
	HQ (NOAEL-Based)	Water ingestion		1.5E-02	1.4E-01
		Sediment ingestion		5.4E-02	1.2E+00
		Aquatic prey ingestion		1.4E+00	2.6E+01
		Total		**1.4E+00**	**2.7E+01**

SMDP 1 at Step 2 of ERAGS

The BTAG considered these results to indicate that inhibition of reproduction was possible in at least some members of the exposed population, but that the fraction of the population that was affected and the degree of impact on the population was difficult to judge from the point estimate calculations. Based on this, a decision was made to conduct a screening level PRA in order to provide some additional information on the magnitude and probability of risk.

Workplan 1

The contractor performing the risk assessment developed a brief workplan that proposed an approach for a screening level PRA. The plan called for a Monte Carlo-based evaluation of variability in exposure and risk among different members of the exposed mammalian omnivore (racoon) population. In brief, all exposure inputs that were treated as constants in the point estimate approach (i.e., were the same for CTE and RME exposure) were also treated as constants in the PRA evaluation. Because water contributed so little to dose or HQ, this pathway was not evaluated in the PRA, but was accounted for by adding in the point estimate values to the PRA results. All variables that are fractions (i.e, may only assume values between zero and one) were modeled as beta distributions, and all other variables were modeled as lognormal. For screening purposes, the parameters for all distributions were selected so that the mean and 95th percentile values of the PDF's matched the corresponding CTE and RME point estimates. The BTAG reviewed this proposed approach and authorized PRA work to begin.

Screening Level PRA Results

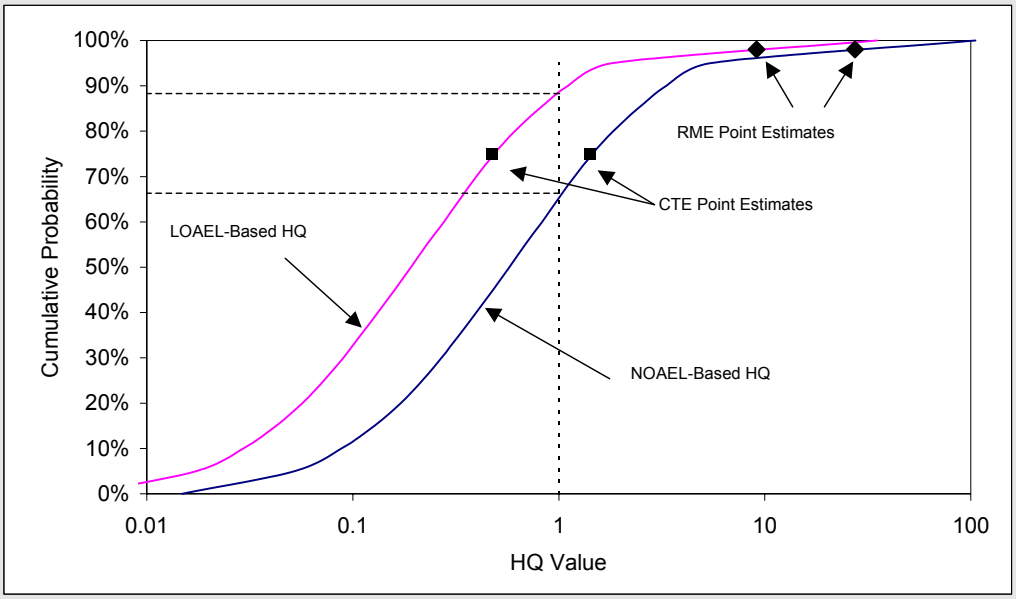
The screening level PRA inputs and the resulting estimates of the variability in HQ are shown in Exhibit 4-11. The CTE and RME point estimates are also shown for comparison. As seen, the PRA distribution of HQ values indicates that about 68% of the individuals in the population are likely to have HQ values below 1E+00, while 32% have HQ values above 1E+00.

Comparison of the CTE point estimates of HQ to the mean HQ values derived by PRA reveals the values are very close. This is expected because both depend on the mean values of the input variables, and the same mean values were used in both sets of calculations. With regard to upper-bound estimates, the RME point estimate values are at the 98th percentile of the PRA HQ distribution, within the target range (90th to 99th) usually considered appropriate. Note, however, that the 98th percentile is about 5-fold higher than the 95th percentile, emphasizing the high sensitivity of the RME HQ values to the precise percentile of the RME.

EXHIBIT 4-11

SCREENING LEVEL PRA CALCULATIONS OF HQ DISTRIBUTION

Data Category	Variable	Units	Screening Level Distribution		
			Type	param 1	param 2
Concentrations	Concentration in water	mg/L	Not evaluated in PRA		
	Concentration in sediment	mg/kg	LN	24	33
	BAF	--	Const	5	
	Concentration in aquatic prey	mg/kg	Calculated		
Intake Rates	Total water intake rate	L/kg-day	Not evaluated in PRA		
	Total food intake rate	kg/kg-day	LN	0.060	0.060
	Fraction of diet that is sed	--	Beta	3.42	110.7
	Fraction of diet that is aquatic prey	--	Beta	6.10	34.6
Area Use Factors	Fraction of total water ingested from lake	--	Not evaluated in PRA		
	Fraction of total diet from the lake	--	Beta	1.20	3.59
TRVs	LOAEL-based TRV	mg/kg-day	Const	0.6	
	NOAEL-based TRV	mg/kg-day	Const	0.2	



TRV Basis	Central Tendency			Upper Bound		
	Mean of PRA	Point Est CTE	Ratio	95th of PRA	Point Est. RME	Ratio
NOAEL	1.44	1.42	0.99	5.4	27.4	5.06
LOAEL	0.48	0.47	0.99	1.80	9.12	5.06

SMDP 2

The BTAG considered these results, and decided that it was very probable that pesticide X was causing an effect in some members of the exposed population, but decided that a final risk management decision would be facilitated by characterizing the distribution of responses (rather than the distribution of HQ values). The BTAG asked the contractor performing the work to develop a proposed approach for characterizing the distribution of responses.

Workplan 2

The contractor obtained a copy of the toxicity report upon which the TRVs were based, and determined that the study did include sufficient dose-response data to support reliable dose-response modeling. The contractor recommended that this be done using EPA's BMDS. The BTAG approved this proposed approach and authorized work to proceed.

PRA Refinement 1

The contractor fit the raw dose-response data (inhibition of reproduction in mink) to a number of alternative models available in BMDS, and found that the dose-response curve could be well characterized by the Hill Equation with nonconstant variance, as follows:

Mean Response at dose d (% decrease in reproduction)=$(100 \times d^{2.5})/(0.9^{2.5} + d^{2.5})$
Std. Dev. in Response at dose d (%)=SQRT$[1.6 \cdot (\text{mean response at dose d})^{1.3}]$

Based on this model, the point estimate LOAEL value (0.6 mg/kg-day) corresponds to an effect level of about 27%, and the NOAEL of 0.2 mg/kg-day corresponds to an effect level of about 2%.

Using this exposure-response model in place of the point-estimate TRV values, the refined PRA predicted a distribution of responses in the exposed population as shown in Exhibit 4-12. As seen, approximately 81% of the population was predicted to experience an effect on reproduction smaller than 10%, while 9% were expected to have a reduction of 10 to 30%, 4% a reduction of 30 to 50%, and 6% a reduction of more than 50%. On average across all members of the exposed population, the predicted reduction in reproductive success was about 9%.

EXHIBIT 4-12

SIMULATED DISTRIBUTION OF RESPONSES

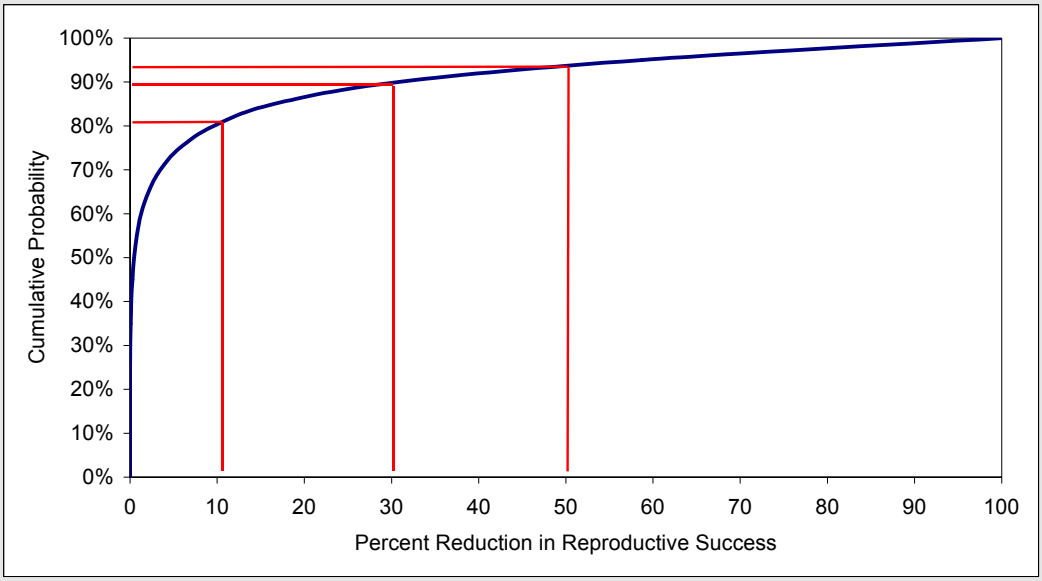
Exposure-Response Model

Resp = Normal(Mean,Stdev)

Mean = a + b*x^n / (x^n + k^n)

Stdev = alpha*mean^rho

x	Total daily intake
a	0
b	100
k	0.9
n	2.5
alpha	1.6
rho	1.3



Percent Reduction	Percent of Population
0-10%	81%
10-30%	9%
30-50%	4%
>50%	6%

SMDP3

The BTAG debated the likely population-level consequences of this predicted distribution of responses in members of the exposed population. After consulting with a field biologist with experience in the population dynamics of mammals such as racoons, the BTAG decided that the distribution of responses in the exposed population would cause a continued stress on the mammalian omnivore community and that reductions in population number were likely over time. Based on this, the risk manager and the BTAG agreed that remedial action was desirable and that a range of alternative clean-up strategies should be investigated. This was performed using the methods described in Chapter 5 (see Exhibit 5-5).

REFERENCES FOR CHAPTER 4

ECOFRAM. 1999a. ECOFRAM Aquatic Report (Draft). Ecological Committee on FIFRA Risk Assessment Methods. Draft report available online at http://www.epa.gov/oppefed1/ecorisk. Report dated May 4.

ECOFRAM. 1999b. ECOFRAM Terrestrial Draft Report. Ecological Committee on FIFRA Risk Assessment Methods. Draft report available online at http://www.epa.gov/oppefed1/ecorisk. Report dated May 10.

U.S. EPA. 1989. *Risk Assessment Guidance for Superfund, Volume II: Environmental Evaluation Manual.* Interim Final. Office of Emergency and Remedial Response. Washington, D.C. EPA/540/1-89/001. March.

U.S. EPA. 1991-present. *Eco Update*. Intermittent Bulletin Series. Office of Emergency and Remedial Response. 1991 to present.

U.S. EPA. 1992a. *Framework for Ecological Risk Assessment.* EPA Risk Assessment Forum. EPA/630/R-92/001. February.

U.S. EPA. 1992b. Policy Memorandum: *Guidance on Risk Characterization for Risk Managers and Risk Assessors* from F. Henry Habicht, Deputy Administrator, February 26.

U.S. EPA. 1994. Memorandum*: Role of the Ecological Risk Assessment in the Baseline Risk Assessment*. Elliott Laws, Assistant Administrator, Office of Solid Waste and Emergency Response. OSWER Directive No. 9285.7-17. August 12.

U.S. EPA. 1995. *EPA Risk Characterization Program*. Memorandum from the Administrator. March 21.

U.S. EPA. 1997a. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments.* Interim Final. Solid Waste and Emergency Response. OSWER Directive No. 9285.7-25. June 5.

U.S. EPA. 1997b. *Guiding Principles for Monte Carlo Analyses*. Risk Assessment Forum. EPA/630/R-97-001.

U.S. EPA. 1998. *Guidelines for Ecological Risk Assessment.* Risk Assessment Forum. U.S. Environmental Protection Agency, Washington DC. EPA/630/R-95/002F. April. Published May 14. *Federal Register* 63(93):26846-26924.

U.S. EPA. 1999. Memorandum: *Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites*. P. Stephen D. Luftig for Larry D. Reed, Office of Emergency and Remedial Response. OSWER Directive No. 9285.7-28. October 7.

U.S. EPA. 2001. *Risk Assessment Guidance for Superfund*: *Volume I. Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments).* Office of Emergency and Remedial Response. Washington, DC. OSWER Directive No. 9285.7-47. December.